Phonologising Articulatory Phonology

Kevin Hind

PhD Dissertation
University of Edinburgh
1997

Abstract

Articulatory Phonology has been criticised as being little more than an enriched theory of phonetics, capable of handling gradient but not categorical phenomena. This thesis is an answer to such criticism, presenting one possible way in which the gestures of Articulatory Phonology can be incorporated into a systematic phonological framework both at the level of the segment and of the syllable.

Segments are created by the combination of gestures in simple head-dependent relationships, where all segments contain one or more heads. A gesture is a head if it dominates the vocal tract, domination being defined in terms of the head's control of neutral articulator settings and of its coordination with other gestures within the segment. Gestural coordination within segments is thus constrained by phonological relationships without recourse to arbitrary distinctions between complete, partial and minimal overlap. These headed structures provide simple accounts of a wide range of segment types such as simple and complex stops, pre- and postnasals, unaspirated and aspirated stops and affricates, as well as a number of common phonological processes such as nasal spreading and lenition. In addition, the use of gestures allows for a description of both gradient and categorical phenomena with a single set of primitives.

Syllable structure is also described in terms of dependency between segments, and constituents are derived from the formal properties of the head-dependency relationships. The structures of Icelandic, Italian and Turkish are examined in detail, with particular attention to the representation of segmental length, preaspiration and epenthesis. Segmental length is represented solely in terms of the phonological relationships between segments and without reference either to external timing units such as moras or x-slots or to manipulation of the stiffness of gestures. It is argued that at least two different types of syllable structures are found, those such as Icelandic where long vowels are V^1 V^2 sequences, and those such as Italian where long vowels are single vowels which are lengthened through their coordination with following segments. In Turkish, which has long vowels of both the Italian and Icelandic type, the two types of long vowel are shown to be phonologically distinct.

Declaration

I declare that this thesis has been composed by myself and that all work contained within it is my own unless otherwise indicated.

Kevin Hind

Acknowledgements

Who to thank? I'll keep it brief, but first of all thanks to both of my supervisors, William Gillies and John Anderson, for keeping faith and showing great patience. I know it took a while, so thanks. Great thanks to Mark Ellison who was willing to hear me wittering on about this and that for a number of years. Finally, and most importantly, thanks to Debbie, as without you it might never have been written.

Phonologising Articulatory Phonology

Contents

Chapter One
The Structures of Articulatory Phonology

1
Introduction

The Sound Pattern of English (henceforth SPE) (Chomsky & Halle 1968) represents phonology as containing strings of linearly ordered segments - the segments themselves being composed of feature sets - devoid of any hierarchical structure or any notions of government or dependency between features or segments. Everything is in a sense created equal. Since its appearance, the number of phonological theories has mushroomed and it would be a fair reflection upon the post-SPE phonological development to say that these theories have moved away from the SPE approach in the same general direction. Current phonological theories are, in contrast to SPE, generally describable as non-linear in that they tend to display at least some hierarchical structure both within and between segments, and have some notions of inequality between units (units being a cover term for features, segments etc.) so that for example feature X may dominate feature Y in some way while the reverse is not true.[1] While these theories are in many ways more complex, as a result they are also more restrictive and more explanatory. Hierarchical structure is present at all levels so that where previously features were grouped together in simple bundles, all features sharing essentially the same type of relationships to each other, segments are now complex structures with their own internal hierarchy, often involving some kind of head-dependent relationships, however this may be expressed. Segments can then themselves combine to form higher-level structures such as syllables or feet.

[1]Rocca (1994) argues that the term non-linear is somewhat of a misnomer, suggesting that the term 'multilinear' be used in its stead. However, while this has some validity, the use of 'non-linear' seems to be well established and I will employ it in this work.

Articulatory Phonology as developed by Catherine Browman & Louis Goldstein (e.g. Browman & Goldstein 1985, 1986, 1988, 1992a) is part of this development. Its chief distinguishing feature is that it builds a theory directly out of the physical movements of speech, thereby claiming that the units of the phonology are also the units of the phonetics, or more simply that there is no distinction between the two domains However, to date there has been a natural focus on the development of the physical generative side of the model, alongwith the analysis of various 'accidental' phenomena such as some cases of segmental deletion and assimilation, phenomena which are to a large extent predictable from the physical nature of the theory. At the same time, a number of abstract units have been developed but there has been as yet little analysis of how these units combine. The result is that while being very much a part of the current movement towards hierarchical phonological models, Articulatory Phonology remains underdeveloped in a number of ways. It is the construction of such abstract structures that I wish to address in this thesis, in particular the question of how Articulatory Phonology can provide a description of segments and how these segments can then combine to form higher structures.

This chapter serves as an introduction to the theory of Articulatory Phonology (henceforth AP), its development and its current form. However it is no more than an outline of the essential points of the theory which are relevant to the discussion in this thesis. For a fuller discussion the reader is directed to the works cited here, in particular to Browman & Goldstein (1989, 1992a) which provide the most detailed discussion to date. The discussion is as follows: In 1.1 I examine the development of gestures in child speech; in 1.2 I describe the development of the theory of Articulatory Phonology, and the role that gestures are claimed to play in phonology and phonetics; finally, in 1.3 I show how gestures can be viewed as occurring within an anatomically bound hierarchy akin to the structures of Feature Geometry.

1.1
The Origin of Gestures

It is generally assumed that speech consists of a number of underlying abstract features at both the phonological and phonetic levels, the number and nature of such features varying from theory to theory. At the same time there is a general consensus that these features cluster together to form a relatively small number of discrete units which we can refer to as segments. At the phonological level these segments may consist of features which can spread, delete, be added, and take part in a number of discrete phonological operations which will have a direct effect on both the ultimate phonological structure and on the phonetic output. Whatever the nature of these phonological processes, phonetically at least we can observe their end results (assuming that there is such a direct connection between the phonology and the phonetics). Observation of the speech - that is, the actual physical production - of adults reveals that it consists of a relatively small set of repeated patterns of movement. For example, a segment such as /p/ may differ in its internal structure from theory to theory, but ultimately every phonology must produce a stable, if variable, pattern of physical movement involving the coordinated action of the lips and the glottis. Noting the existence of these apparently stable patterns of articulatory movement, Browman & Goldstein (1985) hypothesise that speech is built directly from such patterns, the various articulators of the vocal tract coordinating and acting together to form what are referred to as gestures. Importantly, rather than viewing these gestures as being the end product of an underlying phonology, serving solely as the units of phonetic contrast, AP views them as instead forming the basic units of contrast at all levels.

In fact, the existence of gestures would seem to remove the need for distinguishing between levels altogether, as their inherent physical nature would provide both for phonological and phonetic effects. Although there is some abstraction involved in AP's conception of gestures, they nonetheless attempt to model real physical objects and not arbitrary constructs. In addition, the types of physical movements involved in the creation of gestures are not specific to speech but instead form part of an overall model of physical movement, avoiding the need to develop a separate physical model. If this proves to be correct, the relationship between phonology and phonetics will become very much simpler

We can see the origins of gestures in the earliest attempts of children at coordinated physical movement in general, so that the movements specific to speech in fact have their basis in prelinguistic units of action (Studdert-Kennedy 1987). This is clearly visible in the development of babbling in infants. At the earliest stage, beginning around 6 months, children begin to make gross constrictions within the various subsystems of the vocal tract - i.e. oral, velic and glottal - which can on occasion approximate linguistic units present in the surrounding adult speech. The fact that basically the same type of babbling behaviour takes place in deaf as well as non-deaf infants strongly suggests that babbling is at least initially independent of the child having heard adult speech (Oller & Eilers 1988). In the same way, the noises produced by babbling are simply a side effect of the physical movements. At this early stage any resemblance to adult phonological output is a function of adults' tendency to impose their own perception of the structure of their own speech on their children's random utterances.

What in fact seems to occur in babbling is that a number of different articulators are activated to form simple movements within the vocal tract, so that for example the tongue body might be raised to form some type of constriction which may vary from a vowel to a stop. While the actual movement of the various articulators necessary to move the tongue body may be seen as planned to some extent there is no sense of a target or of the constriction being part of any larger phonological structure. At this initial stage children do no more than learn how to coordinate the various muscles and so on of the body in order to produce physical movement, whether this be movement of the limbs or of the vocal organs.

Babbling is initially characterised by a preponderance of tongue body velar articulations, due possibly to the high position of the larynx (Locke 1983), but as children mature and the larynx is lowered, tongue tip and lip gestures increase in frequency and begin to dominate. Overlaying these gross 'consonantal' constrictions on similarly gross vocalic constrictions eventually gives rise to sequences of simple CV syllables which can be roughly identified with the types of syllable structures found in adult speech. These syllables do not form part of a systematic phonology and the transcription of babbling as consisting of discrete syllables is again illusory. These CV syllables arise at the time that repetitive motor functions in general develop, but although they are still prelinguistic structures they are clearly the immediate precursors of real speech (Locke 1986). Once this repetitive stage has been reached children can begin to match their own output with that of the adults around them, harnessing their limited command of articulatory movements to form simple words, often actively encouraged in this by adults, as is evident in the commonality of children's words for parents cross-linguistically e.g. mama, papa. This referential use of gestures is the beginning of phonology, but in order to reach this stage two further skills must be acquired in order to turn a largely random use of articulatory movements into a systematic structure.

Although the vocal tract forms a continuum, languages tend to partition it in discrete ways partly on a universal basis and partly language specifically. For example, a child may produce a single gross tongue tip constriction but may ultimately have to distinguish between e.g. full closure for a stop and critical closure for a fricative, and having opted to form a fricative may then in turn have to distinguish between a number of different locations to distinguish between segments such as /s/, /θ/ and /ʃ/ all using the tongue tip. In other words a child must learn to fine tune its gestures to match the patterns of the target language (Browman & Goldstein 1989). In order to form real speech a child can no longer rely on simple gross movements of the articulators as the language it hears around it makes use of a number of different parameters distinguishing both between the type of constriction employed and the place at which the constriction is made.

Further, a child must learn to coordinate not only the articulators which constitute a given gesture but must also learn to build larger structures, perhaps coordinating more than one gesture to form single segments, and to then coordinate a number of segments to form words. Browman & Goldstein (1989), citing Studdert-Kennedy (1987) present the data in (1) as attempts by a 15 month old girl to say the word *pen* in a single half-hour session. Despite the wide variation in realisations, the girl clearly has some knowledge of the discrete gestures contained within the word but does not as yet know how to coordinate the gestures correctly, nor how to control their temporal and physical extent. The girl has identified the presence of glottal opening, velum lowering, lip closure (or some kind of closure) and so on but has yet to acquire a knowledge as to the correct language particular order necessary to produce the desired output. For example, in [mɑ̃ᵊ] the girl knows to lower the velum but lowers it too early and fails to raise it again. Despite the fact that in many instances the attempts at producing *pen* are far from the desired surface form, phonology may be said to have begun as the child has a clear knowledge that some discrete structure is involved even if her production of those structures is as yet imperfect.

(1)

[mɑ̃ᵊ, ˇʌ̃, dedn, hɪn, ᵐbõ, pʰɪn, tʰṇtʰṇtʰn̦, baʰ, buɑ̦̃]

Gestures in the linguistic sense may be said to be present as soon as a child uses the structures developed in babbling with referential content, though fine tuning of the gestures will continue to occur for some years. Patterns of physical movement which at first are random gradually become more accurate and more stable, but the gestures made by a child when babbling and the gestures made by the same child when producing fluent speech remain essentially the same physical objects. Children aim at the patterns they hear around them in adult speech, identifying the various stable patterns of the target language; this is the primary motivation for the development of AP. If gestures, or at least the beginning of gestures, are present at the earliest point in the development of speech, it is reasonable to assume that they might also be present in fluent adult speech. The strongest assumption that can be made is that gestures are not simply the phonetics produced by a separate series of phonological rules but form themselves the basic units of phonology. This would provide for a maximally simple model of speech where phonology and phonetics were a seamless whole.

If these are valid assumptions and the patterns of speech can be identified with gestures, we must ask both what gestures are and how they are formed into a phonological system. In order to do this we must provide a physical model of gestures, and in the following sections I provide a description of just such a model.

1.2

1.2.1

Modelling Gestures

Gestures, if real physical objects, must be ultimately produced by the coordinated action of a number of articulators, so we must first determine the identity of these articulators and how it is that they combine to produce gestures. (2) below, taken from Browman & Goldstein (1989), is a partial list of the articulators involved.[2] The set of articulators is small, comprising only the lips, jaw, tongue body and tip, velum and glottis. Looked at solely from the production viewpoint we might model speech simply by beginning at the smallest level, coordinating the movements of the articulators without any further hierarchical structure. The internal structure of /pɛn/ in (1) above would then involve initially specifying the muscles involved; the muscles then move the upper and lower lips, the jaw, the tongue body, tongue tip, velum and glottis, specifying the timing, extent, coordination and so forth of each individual articulator. But where do we go from there? By starting from the bottom up we correctly describe the physical movements involved, but in the process create an extremely complex set of rules which nevertheless fail to say anything interesting about speech and also fail to match the observation that speech consists of stable repetitive patterns of movement of the articulators. If on the other hand we begin at the top down, where instead of acting independently of one another the

---

[2]A number of other articulators must also be present, as well as a number of other tract variables, such as those needed to control pitch. The list in (2) however lists all of the variables and articulators currently implemented by Browman & Goldstein.

articulators group together as if they formed single units, we gain a much clearer picture of what actually occurs.[3]

(2)

	Tract Variable	Articulators
LP	Lip protrusion	Upper & lower lips, jaw
LA	Lip aperture	Upper & lower lips, jaw
TTCL	Tongue tip constriction location	Tongue tip, tongue body, jaw
TTCD	Tongue tip constriction degree	Tongue tip, tongue body, jaw
TBCL	Tongue body constriction location	Tongue body, jaw
TBCD	Tongue body constriction degree	Tongue body, jaw
VEL	Velic aperture	Velum
GLO	Glottal aperture	Glottis

The units thus created are known as Vocal Tract Variables. The list of variables currently employed in AP is also given in (2). For every segment we can assume that a number of pieces of information are necessary. To form any oral segment for example we require information as to the correct constriction degree (CD) involved and the constriction location (CL) (velic and glottal segments do not require a specification for CL).[4] In order to form a /b/ we require a CD of complete closure specified by the lip aperture variable, and a CL of bilabial closure specified by lip protrusion, and it is the task of the articulators to achieve these goals. To carry out both of these tasks the upper and lower lips must perform a set of coordinated movements together with the jaw in order to move the lips first together and then apart. In this way we produce the correct articulatory goal which in turn produces the desired acoustic effect.

It is the coordinated movements of the articulators which create the discrete tract variables, and in turn it is the variables which create the gestures. In this way

[3]O course, these units themselves are not, strictly speaking, at the top. The point to be made is that phonology contains units which are themselves composed of smaller features.
[4]Glottal gestures may require a CL specification if we need to refer to the movement of the larynx.

we capture the fact that the articulators do not act independently of each other but instead work together to create discrete physical objects. From the viewpoint of a child hearing the adult speech around him and attempting to match it with his own articulatory input and acoustic output, rather than blindly moving the lips and jaw and only incidentally closing the lips to create a /b/, it is the end result which is important i.e. that the lips be closed. This is what the tract variables reflect. The task is to achieve the correct articulatory goal to produce the correct acoustic effect.

The set of constriction degrees given in (3) can be directly compared with the manner features of other theories, but although they are to some extent quantally based (Stevens 1972), thus containing a degree of audio-acoustic information, they are defined solely in articulatory terms. To return to /b/, the CD of closure involves bringing together the lips until all airflow is blocked at the point of constriction, closure being characteristic of all stops. Of course, if the velum were to be lowered for an /m/ there would be some leakage of the airflow, though not at the point of constriction. The various values given is intended to be sufficient to describe all segment types. Similarly, the constriction locations in (3) can be compared with the place features of other theories, where although the vocal tract in reality is a continuum we can recognise a number of categorical divisions in anatomical and acoustic terms (Ladefoged 1986, Stevens 1989).

(3)

CD : closed critical narrow mid wide

CL : protruded labial dental
alveolar postalveolar palatal
velar uvular pharyngeal

As gestures are real events which have physical and spatial extent, they must also have an inherent temporal extent. However, rather than simply specifying the amount of time taken to produce e.g. /b/, AP instead derives this temporal dimension from the fact that different articulators and variables move at different speeds. The tongue tip, for example, will move a great deal quicker than the tongue body, so if both were to move the same distance the tongue tip gesture would take considerably less time in reaching its goal. In order to capture this distinction AP employs the notion of articulatory stiffness, which can be seen as reflecting the inherent speed and flexibility of the various articulators. The stiffer a gesture is, the less time it takes to achieve its target. Stiffness will vary according to the particular CD, as well as being a function of stress and speaking rate and might thus be seen as conveying primarily gradient information, but in addition it may also be possible to distinguish categorically between glides and their corresponding vowels, e.g. /j/ and /i/ by referring to a difference solely of stiffness (Catford 1977). Ultimately, the knowledge of the distance travelled, together with the speed, allows us to derive the temporal dimension.

These various specifications all combine to form single gestures (4) (those marked with * are currently not implemented[5]). A segment such as /b/ is now specified by a single Lips gesture which combines CD, CL and stiffness. In order to model the actual physical movement of the articulators, the parameters contained within each articulator set in (4) act as the input to a simple task dynamic equation (Saltzman & Kelso 1987; Browman & Goldstein 1989), with the addition of specifications for the mass and so forth of the articulators involved. Task dynamics was originally developed to model the movements of other parts of the body, in particular the movements of the arms in reaching, and its use in the modelling of gestures allows us to see speech as simply part of the overall pattern of movement of the body in general, avoiding the need for a separate model of phonetics specific to speech as noted above. This is just what we would expect if gestures were in fact developed directly from prelinguistic units of action.

[5]Constriction shape may be required to distinguish for example between apical and laminal tongue tip gestures, as well as for laterals.

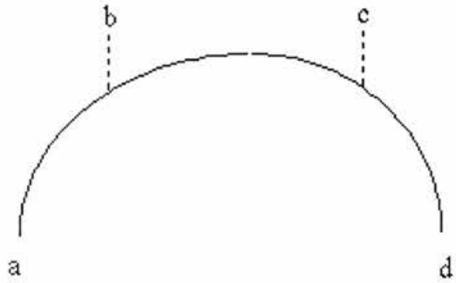
(4)

<u>**ARTICULATOR SET**</u>	<u>**DIMENSIONS**</u>			
LIPS	**(con. degree,**	**con. Location,**		, stiffness)
	LA	LP		
TT	**(con. degree,**	**con. location,**	**con. shape*,**	stiffness)
	TTCD	TTCL		
TB	**(con. degree,**	**con. location,**	**con. shape*,**	stiffness)
	TBCD	TBCL		
TR* (tongue root)	**(con. degree*,**	**con. location*,**		, stiffness*)
VEL	**(con. degree,**			, stiffness)
	VEL			
GLO	**(con. degree,**	**con. location*,**		, stiffness)
	GLO			

The various pieces of information which together form a gesture, such as the information contained within the articulator sets and the numerical values within the task dynamic equations, are known as descriptors. Gestures themselves, however, are single objects despite containing a number of different descriptors, but although they are physically real they are also abstract units in a number of ways. The parameters contained within each articulator set are assumed to be discrete, so that by stating articulatory goals we also assume that the values of the parameters remain constant throughout the period during which the gestures are active. We can thus distinguish between gestures straightforwardly in terms of different values of these discrete parameters, so that /t/ and /s/ are distinguished in terms of their constriction degrees, while /t/ and /t̪/ are distinguished by their constriction locations. While we may describe /t/ as a voiceless alveolar stop and thus implicitly accept that a number of different parameters are necessary for an accurate description, it is clear that we are referring to a single object. The fact that gestures are abstract allows us to view them as phonological objects, but at the same time the physical nature of gestures also makes it possible to describe the most detailed phonetic points.

The task dynamic equation will produce a physical movement which we can represent as a simple sinusoidal wave as in (5), taking again as our example the movement of the lips in forming the closure for /b/. The articulators involved will act together to move the lips towards a closure position, where LP specifies the degree of closure and LS the site at which closure is formed. In this case we have bilabial closure and not for instance labio-dental closure, and the lips form complete closure and not e.g. critical closure which would give /β/. At point (a) the articulators are in a neutral position, the lips lying slightly apart, and at this point the tract variables are activated and the various articulators begin their motion away from their neutral rest position towards their intended target. This target can be said to have been reached at point (b) where the lips are closed. The lips continue to move, compressing and then eventually moving apart, indicating that the CD is a target a little beyond that necessary for closure. At point (c) we can say that closure is no longer maintained and the gesture is no longer recognisably a /b/, although the lips continue to move apart until they are again at rest (d). There is no instant to which we can specifically point and say 'this is a /b/', rather it is the overall pattern created in the vocal tract which we can say is characteristic of /b/.

(5)



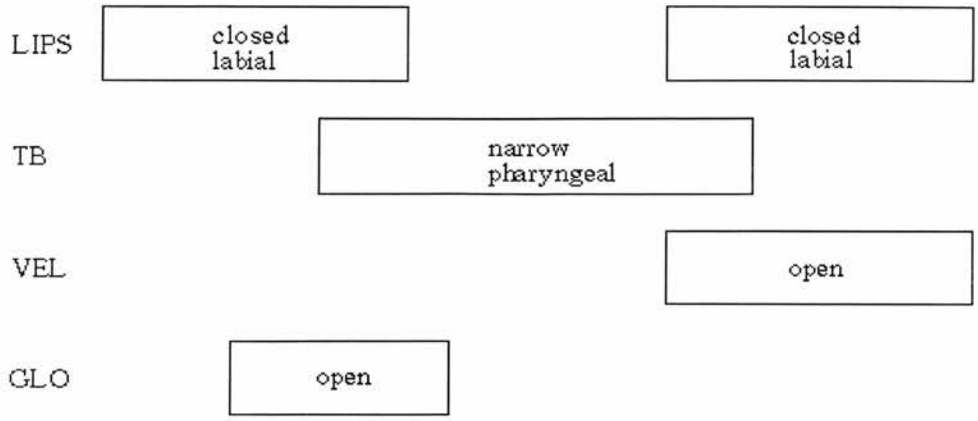
There are two main areas in this conception of gestures which are potentially problematical. The set of constriction degrees, based as they are on purely articulatory grounds, appear to be unable to easily describe processes such as spirantisation as has been noted by Harris (1990). This problem focuses in particular on the seemingly arbitrary change from a CD of [closure] to one of [critical] in e.g. Spanish /b d g/ → [βðɣ], assuming that we can distinguish the CD from the CL as a feature visible to a phonological process. This change is relatively simply described in traditional featural terms as a change in the value of [continuant] from [-cont] to [+cont], but this is of course not an option for AP. Secondly, there is much evidence to support the contention that manner and place features should be independent of each other (e.g. Sagey 1982, Padgett 1991), and if this were to be proved it would argue for a weakening in the relationship between CD and CL. This would be a major blow for AP, as it would seriously undermine the notion of gestures as discrete physical objects. A solution to these problems is needed if AP is to provide full phonological descriptions.

1.2.2
Gestures, Phonology and Phonetics

Having created these real physical objects we must establish how it is that they can be used to create a phonology. The abstract nature of gestures means that they can form abstract phonological relationships with each other, but at the same time these abstract relationships must reflect actual physical coordination. The spatial and temporal dimensions inherent to gestures mean we can illustrate them graphically in the form of a Gestural Score as in (6)[6], which represents 'palm' /pɑm/ (American English). The boxes in (6) indicate roughly the temporal extent of each gesture. The most important point to note is that the nature of gestures means that they can overlap each other both physically and temporally.

[6]The GS in (6) is an idealisation. The exact coordination of the gestures will vary to a greater or lesser degree from utterance to utterance, so that the LIPS and GLO gestures, for example, may show greater overlap in some instances, and lesser overlap in others, resulting in lesser and greater aspiration respectively. Strictly speaking, much of the information contained in the GS, such as the precise coordination between the LIPS and GLO gestures, is predictable and arguably should not be shown. This is discussed in greater detail in chapter 2.

(6)



A number of different physical and abstract relationships are present in the score in (6). While there are three separate segments in /pɑm/, altogether it consists of five gestures. For example, /m/ consists of both a velic lowering gesture and a lips closing gesture, the two gestures coordinating so as to create a single segment. The same is true of the lips and glottal gestures of /p/, where if such coordination were not present we would instead have two separate segments /b/ and /h/. In addition to showing stable patterns of coordination within segments we also find similar stable patterns between segments to create words. In this way the gestures for /p/ overlap the tongue body gesture for /ɑ/ which in turn is overlapped by the gestures for /m/.

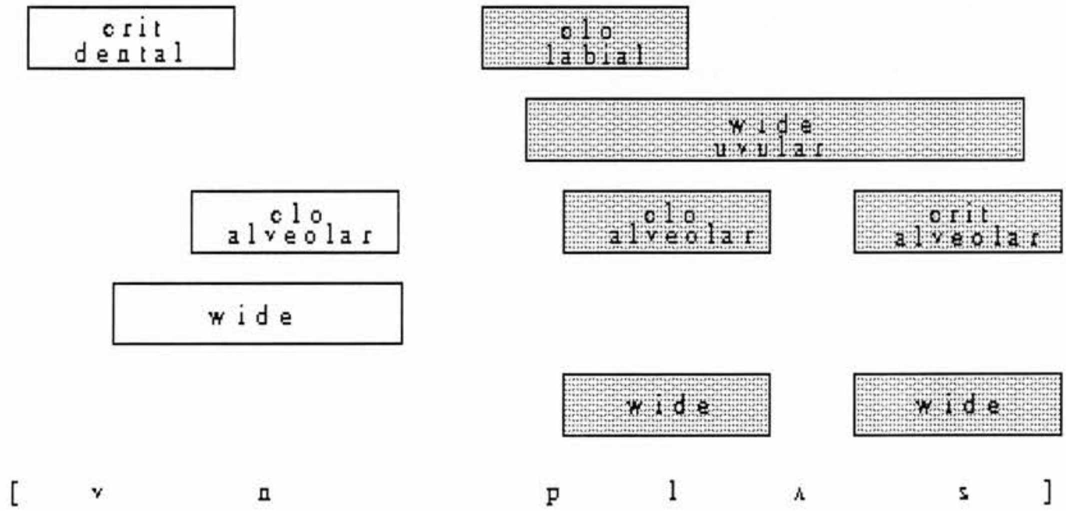
The lexical entry for /pɑm/ will consist of all of these pieces of information: the gestures involved, how the gestures are coordinated to form segments, and how these segments then coordinate to produce the correct surface form. This involves discrete, categorical relationships which are clearly phonological, though the precise nature of the relationships is still an open question, and while there are almost certainly different types of coordination - clearly the lips gesture for /p/ in (6) has a different type of coordination with the glottal gesture than it has with the tongue body gesture for /ɑ/ - all phonological structure is built up from such coordination. Because of the physical nature of gestures, however, the gestural score will also contain information which is generally thought of as belonging to the phonetics. For example, the glottal and lips gestures for /p/ are offset so that the glottal gesture continues to be active after the lips gesture has been released, resulting in a brief period of postaspiration. This must be specified in the gestural score, as must the timing of the achievement of peak opening of the glottis to coincide with the offset of the lips gesture.

The claim of AP is that there is an advantage in capturing both categorical and gradient information without having to place the two in separate domains, a claim which is by no means universally accepted. However, the removal of a boundary between phonetics and phonology would automatically do away with any need to map the one onto the other, a move which should allow for simplification in the modelling of speech. This is not to say that AP is unable to distinguish between categorical and gradient information, merely that while gestures are discrete objects they are nevertheless physical in nature. The lexical entry for 'palm', as noted above, will contain a great deal of both categorical and gradient information, yet even the gradient information such as the exact timing of the achievement of peak glottal opening in /p/ can be considered to show fairly stable forms, at least in citation. In the context of fluent speech however a number of other factors intervene which may affect the actual physical output.

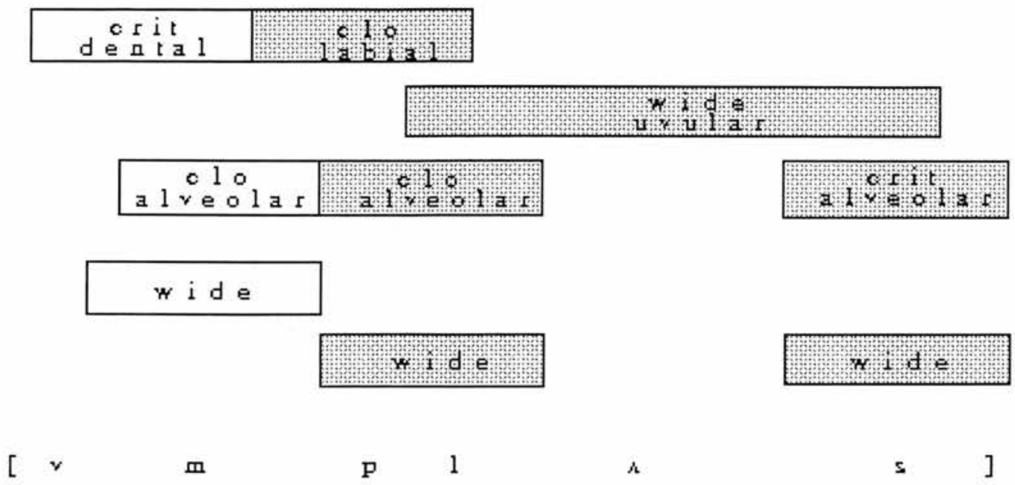
While the patterns of coordination between gestures are presumably stable they may also show a fair degree of variation, so that the velic and lips gestures for /m/, while appearing to be completely coextensive in (6), will not always be so. The velic lowering gesture, being articulatorily independent of the lips gesture, is free to show a great deal of variation as to the timing of its lowering and raising, typically lowering sufficiently early i.e. in advance of the labial constriction, to give rise to some nasalisation of the preceding /ɑ/. This variation will take place within discrete parameters, the exact amount presumably varying on a language particular basis in accordance with the canonical settings of the language in question, and such variation will be present between all gestures, both within and between segments. In citation forms of individual words the variation will have little effect, but in the context of fluent, especially casual, speech with its increased pace, the variation in the physical realisation of gestures, their spatial and temporal extent, can have an extensive effect on the actual and perceived output.

AP makes clear predictions as to what type of variations we would expect in casual speech and the effects thereof. Browman & Goldstein (1987, 1989) discuss the apparent assimilation of /n/ to /m/ in the phrase 'seven plus' in a casual speech context which in more formal speech will have the structure in (7a). This assimilation is generally modelled as a discrete phonological process in which the place features of the /p/, along with constriction specifications, spread to the preceding nasal, possibly deleting in the process any already present place specification. However, instrumental analysis of this type of assimilation (e.g. Barry 1985) reveals that rather than being deleted, the tongue tip gesture for the /n/ remains present in the assimilated forms, arguing against the traditional interpretation.

(7a)



(7b)



The gestures in each word in (7a) show no overlap with the gestures in the other, as would be the case in a careful or emphatic pronunciation. In more casual speech we would expect at least some overlap of the gestures, so that perhaps the gestures for /n/ would be overlapped to some degree by the gestures for the following /p/, as well as by those of the preceding /v/. The tempo is considerably increased in casual speech so that the speaker attempts to make the same gestures but with increasing speed and less time in which to complete them. As a result gestures tend to concertina and show increasing overlap. Increasing the overlap only slightly might have no discernible effect, but as the overlap increases a threshold will eventually be crossed, resulting in (7b) where the tongue tip gesture for /n/ is now completely overlapped by a combination of the preceding lips gesture for /v/ and the following lips gesture for /p/. The velic gesture is now also overlapped by the lips closing gesture, resulting in apparent assimilation but without deleting the tongue tip gesture of the /n/.[7] There is no deletion, no assimilation in the normal, i.e. categorical, sense and no categorical processes. In fact, overlap of only a slightly lesser degree may still allow the /n/ to be perceived as an /n/ and it would be inappropriate to analyse the assimilated forms as showing a discrete categorical change.

---

[7]Assimilation can be be achieved with less overlap than that in (7b) as shown by Byrd (1992) and discussed in Chapter 3.

Overlap is a normal and necessary part of AP and (7b) shows how the natural processes of AP can give rise to apparently categorical changes without the need for any phonological rules. Similar processes can give rise to quite different results. In the phrase 'perfect memory' Browman & Goldstein (1987) note that the final /t/ of 'perfect' is often apparently deleted in casual speech, but that again instrumental investigation shows that the tongue tip gesture for /t/ is not deleted but is simply hidden. As in (7b) the lips gesture for the following /m/ is realised slightly earlier with respect to the preceding tongue tip gesture for /t/. This /t/ is then overlapped by both the following lips gesture and the preceding tongue body gesture for /k/. In this instance however the result is not assimilation but apparent deletion.

Both 'seven plus' and 'perfect memory' involve gestures using different articulator sets, but a different result is predicted when two gestures overlap which employ the same articulator set. In 'ten themes' both the /n/ and /θ/ are tongue tip gestures but with different constriction degrees and constriction locations, and AP predicts that given sufficient overlap the result should be a blending of the parameters of the two.[8] The combination of the two would be expected to give a CL partly alveolar and partly dental, although the exact combined value of CL will vary, and the /n/ is in fact usually transcribed as a dental [n̪].

These three processes, namely complete assimilation, partial assimilation and deletion, are usually modelled as discrete categorical processes particular to casual speech. AP demonstrates instead that they are the natural consequences of a gestural approach. In none of the cases discussed here is there any need for phonological processes as such. The phonology of English and other languages contains information regarding its gestures, their spatial and temporal extent and so forth, and the coordinative relationships which exist between the gestures within the word. Overlapping of gestures will always occur in fluent speech, though it may be

---

[8]The results of this blending are not shown in the GS, though a knowledge of the general apparatus of AP allows it to be inferred.

constrained in a number of ways by the phonology,[9] but the resulting effects in the phrases discussed above are simply blind consequences of the model and need not be specified in any other way.

The gestural score provides a good example of some of the advantages of a gesturally based approach, but it also raises a number of questions about the ability of AP to adequately capture categorical as opposed to gradient phenomena. While AP provides good explanations of a number of gradient processes such as assimilation and deletion, this requires a number of unproven assumptions as to the categorical structures involved. In a word such as 'palm' there are clearly identifiable categorical structures which are nevertheless subject to gradient variation, and we can describe both of these aspects in physical terms, but as yet there has been little attempt to determine the precise nature of the phonological relationships involved (this will be discussed in more detail in the following chapters). In other words, we have yet to determine precisely how the gestures of AP combine to form these apparent categorical structures. If there are segments in AP, what are they and how are they physically realised? Is the use of overlap in the gestural scores too powerful? These and other questions must be answered if AP is to be more than an interesting way of looking at phonetics.

1.3
Gestures Within an Anatomical Hierarchy

While the various descriptors of AP are part of a clear historical development and have analogues in other contemporary theories, they differ from the features of other theories both by being organised into gestures and by operating as part of an overtly physical model. The individual gestures which we have seen so far are to a large degree independent of each other despite in some cases making use of some of the same articulators, e.g. lips and tongue gestures both make use of the jaw yet

---

[9]Browman & Goldstein (1989) note that languages such as Georgian, where all stops must be released in all contexts, can be expected to show differing amounts of overlap from English. Whatever overlap

remain clearly distinct. There are two areas, however, in which the effects of individual gestures on each other are of central importance. As is implicit in the gestural score, gestures can act either alone or in combination to form something equivalent to segments. For example, a glottal opening gesture may stand alone as /h/, or may combine with e.g. a tongue tip closure gesture to form /t/, and in this latter there is clearly some kind of phonological relationship present between the two gestures. There are as many ways of representing such phonological relationships as there are phonological theories, and one of the tasks facing AP is to provide a coherent model of its own. As noted above, a characteristic of post-SPE phonologies is their incorporation of hierarchical structure so that not all features are always 'equal' (e.g. some are dependent on others), and to what extent, if any, AP should incorporate such notions of hierarchical structure is an interesting question. This will be discussed in detail in the following chapters.

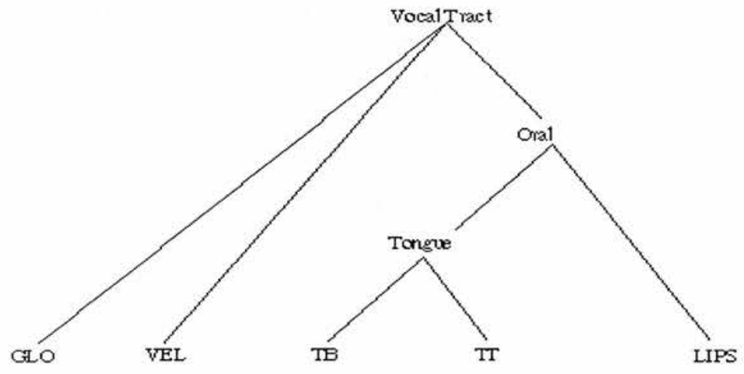
Whatever the relationships which hold between the gestures, their physical nature means that there must be some mechanism for calculating their actual effect on the vocal tract, a problem which becomes increasingly complex as the number of active gestures increases. In order to calculate the combined effects of the various descriptors on each other and on the overall state of the vocal tract, Browman & Goldstein suggest the incorporation of an articulatory geometry. Such a geometry is a characteristic of Feature Geometry (FG), where in contrast to many other theories, the various features are linked together in a fixed hierarchy. The actual structure of the hierarchy is a matter of some debate, resulting mainly from the premise that the geometry, while in part derived from anatomical considerations, must be determined primarily on the basis of phonological patterning. We will examine some of the claims of Feature Geometry in subsequent chapters, looking in particular at some of the areas where AP is claimed to offer a superior approach. However, while AP adopts a geometry which is superficially similar to that of Feature Geometry, it bases it solely upon the actual geometry of the vocal tract.

---

did occur in casual speech in such languages would be restricted by the requirement that the stops have an audible release. Similar constraints will operate in other languages.

Gestures take place within a physically real vocal tract, and the interactions of gestures must be constrained by it. It therefore seems a reasonable supposition that as relationships between gestures are constrained by the physical reality of the vocal tract, any hierarchy we propose for AP should be anatomically based, phonological patterns and phonological rules deriving from the already existing hierarchy rather than vice versa. The geometry in (8) below represents an initial attempt at such a hierarchy, based solely on the physical relationships between the articulator sets in (4). The gestures function as the terminal nodes of the tree and a number of connections are made between them. The tongue tip and tongue body gesture form a single tongue node on these anatomical grounds, and together with the lips gestures they form a class of oral gestures. Each of the gestures within this larger class are characterised by their common use of the jaw as an articulator. Combining these with the remaining velic and glottal gestures gives us an overall geometry similar in many ways to those proposed for FG.

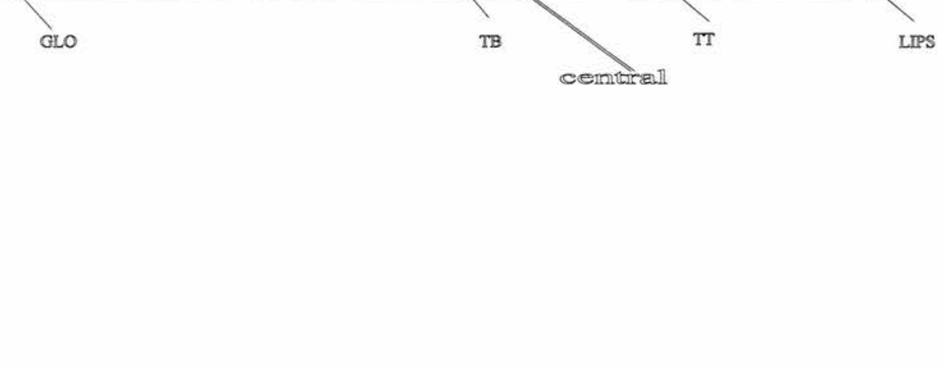
(8)



We can now create natural classes on purely anatomical grounds, matching the division between oral, velic and glottal gestures found in FG. For example, certain coronal consonants such as palatals may involve both the tongue tip and tongue body, so may best be classified at the tongue level (Keating 1988), and we can continue to group other parts of the tree together, such as the lips and tongue, to provide further contrasts. However, while this geometry may be useful on a very general classificatory level, it provides us with next to no information as to the actual physical, audio-acoustic effects of gestures. It is not enough simply to note the existence of gestures. We must also be able to say what these gestures do, in particular how they affect the airflow through the vocal tract and ultimately produce sound. In order to do this we must create a hierarchy based not only on the gestures but on the environment in which the gestures take place, i.e. the vocal tract. Tube geometry provides just such a model.

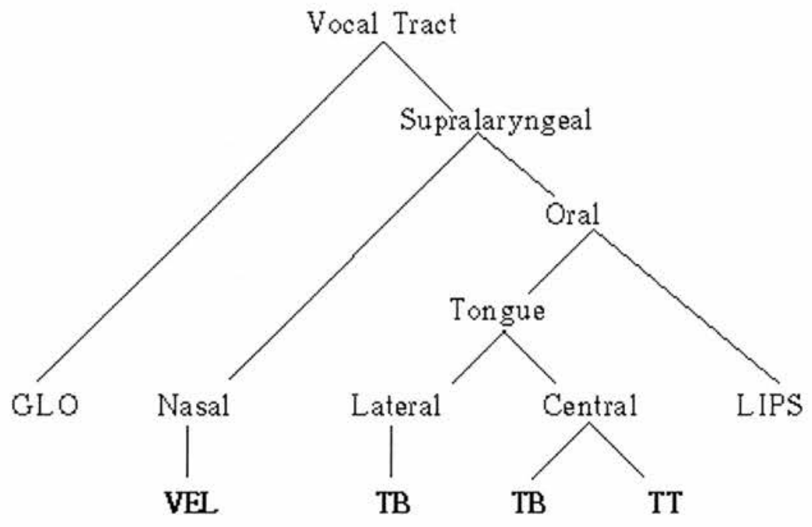
Tube geometry views the vocal tract as a set of interconnecting tubes organised by the anatomy as in (9). The vocal tract is envisaged as containing three main tubes - a nasal tube created by the lowering and raising of the velum, a central tongue tube, and a lateral tube[10] - and at either end these tubes are terminated by glottal and lips gestures. The various articulators move within these tubes, i.e. the individual gestures take place within the tubes, and the overall state of the vocal tract is a result of the constrictions made within the tubes by the gestures. From this tube geometry we can construct the hierarchy in (10) which contains a more complex set of interconnecting tubes which ultimately converge at a single point at the top of tree from which we can read the overall audio-acoustic output.

(9)



[10]The tongue body is involved in creating two separate tubes. By narrowing the sides of the tongue to allow air to flow past it we create a separate lateral tube.

(10)



The bottom of the tree in (10) represents the gestures forming the three basic tubes and their terminators. The upper levels of the tree are created by combining the simple tubes with each other and with the terminations of the tubes, i.e. the glottal and lips gestures, to form more complex tubes, The tongue tip and tongue body gestures combine to form a central tube, and this then combines with the lateral tube to form a still more complex tongue tube. This tongue tube itself is then terminated by the lips to create a combined oral tube, the oral tube then combining with the nasal tube to form a supralaryngeal tube. The top of the tree is formed by the intersection of this complex supralaryngeal tube with the termination created by the glottis to create a single tube dominating the entire vocal tract. In other words the entire vocal tract can now be viewed as a single tube built up from a number of smaller tubes.

Given this hierarchy, the constriction degree within each tube at any one instant can be calculated from the combination of the various gestures active within it. Each tube will have its own neutral setting so that there is some specification within the tube even when no active constriction is being made. For example, when forming a /t/ the lips play no active part yet we must ensure that they do not form some constriction which would affect the intended output. In fact, during the production of a /t/ the lips typically lie open and so we specify this position as the neutral setting. The full list of neutral settings is given in (11) (CD = constriction degree). As noted earlier, although the lips play no active part in creating a /t/, in a word such as 'too' lip rounding is required for the vowel and may actually begin during the consonant as it does not conflict with any requirement for lip movement in /t/. Similarly, in 'you' lip rounding is again required for the vowel, but lip spreading is also required for the consonant; in this case the two values for the lips will blend.

(11)

Nasal	[CD] = clo (i.e. closed)
Lateral	[CD] = Central [CD]
Central	[CD] = open
LIPS	[CD] = open
GLO	[CD] = crit (i.e. critical)

The tubes themselves form two different kinds of relationships. The tongue tube is formed from the combination of the central and lateral tubes, and as is clear in (9) these two tubes are connected in parallel. Airflow through the tubes will follow the path of least resistance, and thus when two tubes are joined in parallel the combined CD will be that of the least constricted tube, the minimum CD.[11] When joined in series the reverse is true, that is the CD is that of the narrowest tube, the maximum CD. We can see this in the oral tube where if the tongue has an open or critical CD but the lips are closed, the CD at the oral level will be closed, i.e. there will be no airflow.[12] The oral tube then forms the more complex supralaryngeal tube by joining in parallel with the nasal tube. If the CD of the oral node were to remain closed but the CD of the nasal tube were open, i.e. the velum was lowered, then the CD of the supralaryngeal tube would also be open, allowing airflow through the nasal passage. The full pattern of percolation of CD through the hierarchy is given in (12).

[11]I follow here the usage of Bird (1990) rather than that of Browman & Goldstein (1989).
[12]The lips and glottal terminations act as if they created tubes joined in series.

(12)

Nasal	[CD] =	VEL [CD]
Lateral	[CD] =	TB [CD, CS = narrowed]
Central	[CD] =	MAX (TT [CD], TB [CD, CS = normal])
Tongue	[CD] =	MIN (Central [CD], Lateral [CD])
Oral	[CD] =	MAX (Tongue [CD], LIPS [CD])
Supra	[CD] =	MIN (Oral [CD], Nasal [CD])

The principles in (12) do not account for the CD at the top of the tree, the vocal tract level. Here the combined values of the gestures dominated by the supralaryngeal tube join both with the output of the glottal tube and with the pulmonic initiation to create the overall aerodynamic and acoustic value of the various gestures and default settings. In other words, we are no longer concerned with articulation alone but also with its consequences. (13) gives an approximation of both the types of articulatory constrictions involved and of their non-articulatory interpretation. Occlusion in either the glottal or Supra tube will naturally lead to closure of the vocal tract as a whole, one of the consequences of which will be a blocking of the airflow with all that entails. The Supra tube may instead have a critical or open CD, and in this case the role of the glottis will determine the overall value of the vocal tract. If the glottis is closed we will have occlusion no matter what the state of the remainder of the vocal tract; if it is not, the result will be noise if the Supra tube has a CD of [crit], the type of noise depending on the state of the glottis, or resonance if the Supra tube has a CD of [open]. The vocal tract then can have a value of occlusion, noise or resonance, the exact nature of each of these values being determined by the values of the constrictions in the tubes beneath it.

(13)

VT outputs

occlusion:	no airflow through VT; silence or low amplitude voicing	
noise:	turbulent airflow	
resonance:	laminar airflow with voicing; formant structure	
VT (CD)	= occlusion	/ Supra [CLO] or GLO [CLO]
	= resonance	/ Supra [OPEN] AND GLO [CRIT]
	= noise	/ otherwise

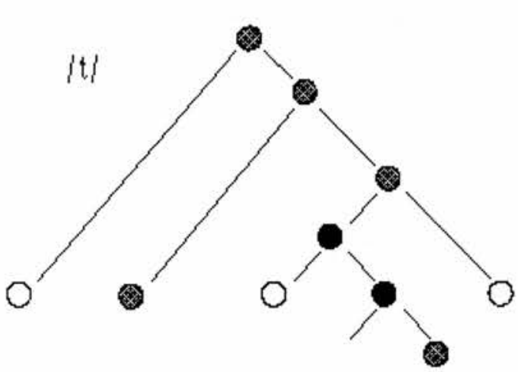
The resulting hierarchy, while similar to those proposed for FG in many ways, differs from them in three important respects. First of all, the hierarchy in (10) is purely anatomical and thus universal. There can be no instances where languages form different hierarchies to capture different phonological distinctions as the anatomy of the vocal tracts of individual speakers is to all intents and purposes identical from speaker to speaker. Instead different phonological patterns arise from different interpretations of the patterns of constriction formed within this anatomy. Secondly, the vocal tract level at the top of the tree is not an abstract node but instead reflects the overall condition of the vocal tract at any one instant. Thirdly, and most significantly, every level of the tree without exception has its own CD, whether created by an active constriction within the tube or resulting from a default setting.

The presence of CD at all levels of the hierarchy simply reflects the physical realities. During speech the various articulatory movements form a set of overlapping constrictions which necessarily means that the state of the vocal tract is constantly changing. No matter which phonological theory we choose to follow, it remains true that during speech every part of the vocal tract has some kind of constriction, constriction here being a cover term for anything from complete occlusion to complete openness. The percolation principles then simply act as a system of checking upon the overall state of the vocal tract at any instant.

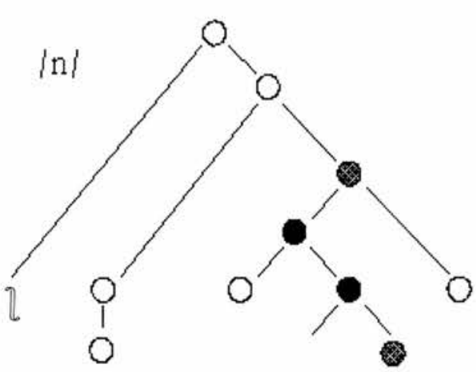
Natural classes can be now be distinguished by differences in the CD at various levels of the hierarchy as in (14) (filled circles indicate a CD of closure, empty circles one of open; a waved line indicates a critical constriction). The voiceless stop /t/ contains two gestures, a tongue tip closure gesture and a glottal opening gesture, and this is a sufficient description. The vocal tract hierarchy created by these gestures in (14a) shows the closure from the tongue tip gesture dominating so that the overall value of the vocal tract is one of occlusion accompanied by voicelessness.[13] (14b) shows /n/ which also contains a tongue tip closure gesture, but which has an additional velum lowering gesture. Both consonants can be seen to form a natural class at the oral level, which for both has a CD of closure, and the same would be true if we replaced /n/ with e.g. /m/ or /t/ with /k/. At higher levels however the stop and the nasal are clearly distinguished as the lowered velum creates an open supralaryngeal tube. The nasal forms natural classes at this level both with any segment which does not show closure at the Supra level (anything other than a stop in fact), or more closely with a segment which has an open Supra tube, such as a sonorant or vowel.

(14)

a.



b.



Reading off natural classes from the hierarchies in (14) of course begs the question as to the phonological status of the vocal tract hierarchy in AP. While in FG the top of the tree is occupied by an abstract root node, in AP the tree represents a single complex tube, the top of the tree simply providing a way of calculating the overall state of the vocal tract. By using the tree to distinguish between natural

[13]Such domination is not, strictly speaking, determinable solely from (14a). This is of course only

classes rather than relying solely on the gestures - /t/ and /n/ both have closure gestures - the implication is that rather than simply being a way of calculating the overall state of the vocal tract at any one time, the vocal tract hierarchy might itself be a phonological object on which various operations might be performed, such as spreading, deletion and so on. Deletion of entire gestures would certainly have an effect on the output. If, for example, we were to remove the tongue tip gesture from /t/ we would be left with only the glottal opening gesture i.e. /h/. Removal of either CD or CL but not both is however clearly not possible, at least under normal circumstances. In Feature Geometry, the separation of manner and place allows the one to be altered, or even deleted, without affecting the other. If similar operations take place on the trees in (14), however, it seems as if the lowest level visible to such rules would be the gesture. In other words, the rules would not have access to either the CD or CL as separate specifications, but only both together as part of the gestures.

This is part of a more general problem in AP. As we have already seen, the distinction between categorical and gradient information is not clear cut. A number of categorical relationships are assumed to exist, as exemplified in the gestural score, and categorical phenomena should therefore be easily distinguishable from gradient, but the necessary phonological relationships between gestures have yet to be identified. Given an onset sequence such as /tr/ we are able to say a number of interesting things. We can distinguish /t/ from e.g. /d/ by the presence of a glottal opening gesture in the former, so that presence versus absence of gestures is sufficient to provide categorical contrasts. Crucially we assume that the relationship between the gestures in /t/, and between these gestures and those for /r/, are somehow different, as reflected in the physical coordination between them, but precisely what these relationships are is still to be determined. This is made all the more difficult by the fact that typically in English the glottal opening gesture may completely overlap the tongue tip gestures for both the /t/ and the /r/. We need to provide a non-arbitrary

---

possible if we have access to other apparatus of the theory.

dividing line between what is gradient and what is categorical whilst describing both in terms of the same units.

It is these problems and others that the remainder of the thesis addresses. In chapter 2 I look at the nature of segments in order to identify what type of structures need to be incorporated into AP. In particular I discuss the nature of categorical and gradient information to determine how each can be incorporated into the phonology. In chapter 3 I present a theory of segment structure, showing how simple head - dependent relationships can account for all segment types, and how we can distinguish between categorical and gradient information whilst showing how gradient phenomena can be vital to a full phonological explanation. In chapter 4 I examine the types of relationships which can exist between segments in AP, both physical and phonological, building a theory of syllable structure again from simple head - dependent relationships. In chapter 5 I examine some of the consequences of the view of phonology put forward in the previous chapters, along with some continuing problems for AP intrinsic to the gestural approach. Finally, in chapter 6 I give a brief summary of the arguments developed.

CHAPTER 2
CATEGORICAL VS GRADIENT

2
Introduction

While the exact placement of the boundary is often unclear, the existence of the two separate but related fields of phonology and phonetics has long been recognised. At the simplest level, phonology is what we think, phonetics is what we speak. The physical act of speech involves real physical movement of the various articulators of the vocal tract, and to this extent we are bound by the physical geometry of our vocal systems as to the range of sounds we can produce. A different set of articulators in a different vocal tract would undoubtedly lead to a different set of sounds being produced. However, speech is more than just the sounds produced, and it is generally assumed that phonology is an abstraction over the phonetics, as stated by Trubetzkoy (1969) (quoted in Coleman 1992) :

> The data for the study of the articulatory as well as the acoustic aspect of speech sounds can only be gathered from concrete speech events. In contrast, the linguistic values of sounds to be examined by phonology are abstract in nature. They are above all *relations*, oppositions, etc., quite intangible things, which can be neither perceived nor studied with the aid of the sense of hearing or touch.

Thus phonologists have set themselves the task of analysing languages with a view to identifying Trubetzkoy's relations, oppositions and so on. We know that languages can and do differ as to the details of their phonologies, but the assumption is that we should be able to develop a set of principles, a methodology, which will allow us to determine what these relations and oppositions are, for any language. In addition, despite the seemingly large differences which exist between individual languages, there is also an assumption that fundamentally all languages operate within the same set of restrictions, enabling us to establish a theory of phonology which will be universally applicable. By claiming that there exists an underlying, abstract system containing elements which may, but do not necessarily, form one-to-one relationships with the elements of phonetics, phonetic 'facts' are irrelevant to the extent that identical phonetic data in different languages could be organised in different ways in the respective phonologies. At all times it is the underlying system which is important.

This belief is compatible with very different interpretations of the relation between phonology and phonetics. On the one hand there is the widespread (and perhaps dominant) view that the two form a continuum, and that they can therefore be expressed using the same formal language. This is the position of theories as diverse as Dependency Phonology (Anderson and Ewen 1987) and Feature Geometry (e.g. Clements 1985), where various elements, or distinctive features, are held to be universally present in phonology, and that these features are directly mapped into the phonetics. On the other hand, we can take the position of Trubetzkoy quite literally, and assume that there is no formal resemblance between the two domains, the features of the phonology being 'contentless' in the sense of Coleman (1992), i.e. not inherently meaningful. To this extent, there is no direct, one to one mapping from phonology to phonetics. However, despite these differences, the one thing on which these theories all agree is that phonology and phonetics are discrete objects, that each can be examined without reference to the other.

It is this boundary between phonology and phonetics, between the system and the sounds it produces, that is 'blurred' in AP. Rather than use the same language to describe the two domains, with some system to map one on to the other, AP, on the broadest interpretation, simply deletes the boundary. Coleman (1992) has claimed that on this interpretation phonology becomes no more than a partial phonetic description, but we might equally claim that phonetics becomes a fuller phonological description. Whichever view we choose, we no longer have two domains, only one.

However, as it stands AP is inadequate in that while it has successfully described a number of otherwise puzzling casual speech phenomena, as noted in chapter 1, it has so far failed to provide equally convincing accounts of anything outside of this restricted range. AP has been extensively criticised because of this (e.g. Padgett 1991; Clements 1992; Kingston & Cohen 1992), and as yet the theory has not been enriched in order to answer these criticisms. In this chapter, I will examine the mechanisms employed by AP, showing how they fail at present to adequately describe the data given, and what problems the data bring to phonology in general. In **2.1** I discuss the question of whether categorical and/or gradient information should be present in phonology; in **2.2** I examine the nature of the categorical structures of AP, dealing in particular with the representation of aspiration in both AP and Feature Geometry; in **2.3** I discuss the existing proposals for extending the role of categorical structures in AP; in **2.4** I discuss the representation of complex and contour segments in AP and FG, showing how neither adequately describes the data; finally, in **2.5** I examine the proposals of Bird (1990) who suggests a new analysis of the role of the Vocal Tract Hierarchy in AP.

2.1
The Nature of Phonology

One of the oft-cited characteristics of phonetics is that it is gradient, whereas phonology is categorical, and indeed Padgett (1991) goes so far as to claim that this fundamental difference between the two domains demands that different mechanisms be used to describe each. To date, many have interpreted AP as covering a restricted domain, one which coincides with that of phonetics rather than that of phonology. This has been possible largely thanks to AP's limited success with describing discrete, categorical effects together with its relative success in describing gradient phenomena, thus apparently showing itself unable to capture anything other than a limited degree of categorical information. A consequence of this view of AP is that certain phenomena have been moved from the phonological domain into the phonetic, where it is thought that a gestural approach might provide a more elegant analysis, with double benefits in that not only are these phenomena explained, but they are also removed from the remit of phonology. The 'hiding' of /t/ in 'perfec(t) memory' seen in Chapter 1 is one such area where it might appear as if we can simply remove a problem from the phonology and pass it on to some (often ill-defined) phonetic domain.

The question of whether AP's analysis of such apparent deletions as in 'perfec(t) memory' should be treated as phonological is clearly an important one, since for a theory as phonetically based as AP it is vital to know which data to consider, and how widely the net should be cast. Padgett (1991) discusses Clements' (1985) treatment of coronal assimilation in English in this light. The data in (1) below show how /t d l n/ all assimilate to varying extents to following coronals. As Padgett notes, the fact that the assimilation applies across words, is non-structure preserving and is 'more or less automatic' suggests that 'phonetic features like gestural overlap' may be at work, rather than a 'strictly' phonological rule.

(1)

a.

eighth	(t,d,n,l ---> interdental)
hundredth	
tenth	
twelfth	

b.

eight chairs	(t,d,n,l ---> postalveolar)
red shoes	
inch	
pilchard	

c.

tree	(t,d,n,l ---> retroflex)
red roses	
enrol	
all right	

The data here would indeed be simply analysed as resulting from blending, with two consecutive coronal articulations, using the same articulator, finding some middle ground in terms of [place]. As Padgett points out, instead of fully sharing [place], as a FG account would predict, we find compromise, so that e.g. instead of being fully retroflex, the /t/ in tree appears to 'glide' from its alveolar position through to postalveolar.

Rather than taking this as evidence for expanding the power of the phonology to take such phenomena into account, Padgett assumes that here we have an argument that this process should be taken out of the phonological domain, as we have not a discrete categorical shift but a gradient one. In other words, it seems as if Clements has cast the phonological net a little too widely by providing a phonological description of a phonetic process. Padgett's position is important because to date his proposals regarding the structure of the FG featural tree are the closest, both in content and in spirit, to AP. However, it is clear that Padgett regards AP as having at best an interpretative function, as a mediator between the abstract representations of his own theory and the phonetics. He insists on the categorical to the exclusion of the gradient, rather than incorporating both, and makes his position quite clear : "....a theory of phonology must capture the categorical (*as opposed to* gradient and continuous) nature of phonological rules" (Padgett 1991, p38; my italics). In other words, any theory which includes gradient and continuous information is by definition not a phonological theory.

The alternative to Padgett's approach is to devise a phonology which incorporates both categorical *and* gradient information into its structure. This is the approach taken by AP, with its emphasis on the incorporation of real time, physical movements and hierarchical structure, but so far, while there has been a fair amount of success in tackling casual speech phenomena which appear to involve gradient processes, there has been far less success in describing categorical processes.

The criticism which Padgett levels at AP is thus quite straightforward: phonology deals with categorical processes operating upon categorical structures, but AP contains either not enough or not the right kind of categorical structures, and as a result appears to be able only to account for gradient processes. Therefore, it does not constitute a phonological theory. However, while the primacy of the role of categorical structures and processes would be agreed on by most, if not all, current phonological theories, there is far less agreement even within individual theories as to what these processes and structures actually are, as illustrated by Padgett's gradient reinterpretation of Clements' putative categorical assimilation process, and the various reinterpretations by AP seen in chapter 1. This is true both in terms of defining the boundary between what is categorical and what is not, and in determining what these categorical structures and processes are. Regardless of this, the importance of the categorical is clear, but what advantages, if any, does its incorporation give us?

Looking first at categorical processes, let us assume that Clements is correct in his interpretation of (1) above, so that we have a process of the type X→Y \ _ Z i.e. a unit X becomes a different unit Y in the environment Z. This is clearly a categorical change, with one unit becoming another. So, if coronal assimilation in (1) were an active phonological process in English, the alveolar nasal /n/ in 'tenth' would be changed to the dental nasal /n̪/ by a combination of feature spreading and feature deletion. There is no halfway house possible; either assimilation takes place or it does not. This is one important aspect of what Padgett refers to, the ability to make discrete all or nothing changes to the structures within the phonology, and similar processes are envisaged within AP, e.g. deletion of oral gestures in Maya (Browman & Goldstein 1989). However, although such categorical processes are as yet scarce within AP, the reason for their scarcity lies mainly in the fact that developments within the theory have been concentrated elsewhere, and there is nothing within AP to suggest that similar types of processes to those found in FG and other theories could not be developed. That said, without categorical structures there can be no categorical processes, and this is an area of very real concern for AP.

Wherever the boundary between categorical and gradient may lay, there is general agreement that from the stream of speech we can identify recurrent patterns, relations, groupings, classes, or whatever term we wish to choose. From these patterns are formed sets of basic phonological units, and these units are combined to form still higher units, which in turn may form still higher units and so on. Yet although all theories draw from the same pool of information they disagree in many respects as to the nature of these units, as well as to the ways in which they should be combined. There is one area, though, in which we can find a broad base of agreement, and that is the segment.

Although the segment is no longer the most basic unit of phonology, it still plays a central role in most current theories. The segment would appear to be the archetypal example of a phonological unit , but as with all such phonological patterns it is difficult if not impossible to devise a description of it which would be agreed on by all theories, and even within individual theories this remains a problem to some extent. This lack of common ground, though, lies less in any profound disagreement as to the basic nature of segments than in the difficulties of translating one theory-specific formulation into another theory. Given an utterance /pis/, most if not all theories would agree that it contains three separate units or segments - /p/, /i/ and /s/ - which combine to form a more complex single unit, a word. This division of the utterance into three discrete units is of course reflected in the phonetic transcription, and the description of each as being a single units holds good despite their each containing a number of different phonetic parameters, precisely because these different parameters consistently pattern together. What matters is that within a phonological system, whatever elements comprise e.g. /p/ can behave as if together they formed a single unit in contrast to other possible combinations, so that every segment may be said to individually form a natural class containing a single member.

Similarly, even though they might appear to be more complex structures than /p/, both /p^h/ and /gb/ would also be generally agreed to be single segments based on their behaviour within phonological systems. Thus, we encode e.g. /gb/ as a single segment and not as two because of the fact that we can find a great deal of both phonological and phonetic support for claiming that it behaves in a way comparable to segments such as /p/ rather than as a combination of two separate segments /g/ and /b/.

There is less agreement about some other putative segment types, such as affricates and prenasals. As we shall see in greater detail in chapter 3, the question of whether the likes of prenasals constitute single segments or sequences of some kind depends both upon the limitations or otherwise of the particular theory employed and on the way in which such sequences pattern within a particular language. Despite limitations of this kind, it seems clear that although the assimilation in (1) targets a single feature, [coronal], the ultimate effect is to change one discrete segment into another. Without some means of encoding this notion of segment we would not be able to express this fact.

Segments may be said to form phonological events in the sense of Bird & Klein (1990), where an event consists of an interval together with a property, so that for e.g. /p/ there is an interval during which it is true that there is bilabial closure accompanied by voicelessness. However, while it is true to describe both /p/ and /p^h/ as being single units and therefore phonological events, it is nevertheless also true that just as these units can combine to form higher units such as syllables, so they are themselves divisible into a number of constituent parts. In other words, events may themselves be composed of events, which may be gestures or, in the case of FG, features. These features of FG serve as the building blocks of segments, and are themselves categorical units designed to represent recurrent patterns. Unlike gestures, the featural system is based on a mixture of articulatory and acoustic data so that e.g. [dorsal] represents a set of related and well defined articulatory patterns while [sonorant] similarly represents a set of related acoustic patterns. Features (or autosegments) are autonomous in the sense that they are discrete units and are free to take part in phonological processes independently of other features. However, crucial to FG, and characteristic of it, is the way in which features are also linked to each other in a fixed hierarchy.[1]

[1] See Piggott (1992) for arguments that rather than there being a single fixed hierarchy, there may instead be a restricted set of possible hierarchies.

This hierarchy reflects the fact that certain features consistently pattern together in phonological processes while other features do not. For example, FG recognises the features [dorsal] and [labial], and each of these can act independently of the other. Given a complex segment such as /gb/ containing both [dorsal] and [labial] nodes, an adjacent segment could theoretically assimilate to either one of these nodes, so that a preceding nasal assimilating to [dorsal] would result in [ŋ] while assimilation to [labial] would give [m]. What typically happens of course is that a nasal will assimilate to both nodes, resulting in [ŋm], suggesting that in some way [dorsal] and [labial] can act as a single unit. This behaviour is captured by making both nodes, together with [coronal], dependent upon the [place] node. Now, instead of having to specify that the nasal assimilates separately to both [dorsal] and [labial] the rule is more simply expressed by specifying that assimilation targets the [place] node, automatically involving assimilation to both [dorsal] and [labial].

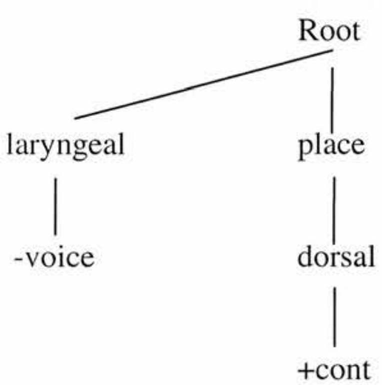
These kinds of relationships, designed to express in an optimal fashion the patterns which we find in speech, are repeated elsewhere in the hierarchy, and are expressed in terms of asymmetric relations between features such as A dominates B, or B is dependent on A. Thus, [dorsal] is dependent upon [place], i.e. [place] > [dorsal], and we can say that a phonological relationship exists between the two such that if [dorsal] then [place].[2] These phonological relationships have many important consequences, e.g. a deletion process which targeted the [place] node of /gb/, given [place] > [dorsal] and [place] > [labial], would lead to deletion of both its dependants, resulting in deletion of all specification for place of articulation. Alternatively, deletion might instead target e.g. [dorsal] without affecting either [labial] or [place], resulting in /b/.

There is, then, a series of phonological relationships between features in FG of the type [place] > [dorsal], and in many ways the groupings created by these relationships are directly comparable with the gestures of AP. For example, the

[2] [dorsal] is said to be the 'daughter' of [place]. Asymmetric relations also hold between the various dependants of [place], which are referred to as 'sisters'. Sisters show the relation of linear ordering, some consequences of which are discussed below in section 2.4.

sequence [place] > [dorsal] > [+continuant] (assuming the geometry of Padgett 1991) for /x/ contains much the same information as does the gestural specification TB : **cri** of AP. However, for a fuller specification of /x/ in FG we must also provide information as to the role of the glottis, in this case [laryngeal] > [-voice], as in (2). We now have two sets of information, one regarding [place] and its dependants, the other regarding [laryngeal] and its dependant. While both of these groupings can be accessed individually by phonological processes, they also act as if together they formed a single larger grouping, a single unit; in this case, of course, the unit that they form is /x/. This grouping is expressed in the same way as are all other groupings, that is by making all features which are part of the grouping dependent, whether directly or indirectly, on a single feature, the root node. As a result, just as [dorsal] is dependent on [place], so [place] is dependent on the root node, the phonological relationships being the same in both cases, thereby avoiding the necessity of introducing any new type of relationship between the features.

(2)



The Root node's role thus corresponds directly to the notion of phoneme or segment (Sagey 1986), so that despite the relative freedom of the features in (2), their binding together by the root node ensures that they can together be regarded as a single categorical object. Thus, just as a phonological process targeting [place] would automatically also target the dependants of [place], so processes targeting the root node would automatically target all of its dependants. For /x/ this would include not only the [place] node and its dependants, but also e.g. the [laryngeal] node and its dependant [-voice]. The geometry in therefore (2) directly expresses this fact.

This is where AP fails as a phonological theory. As we shall see, there are no comparable structures in AP, with the result that it is impossible to know how AP could resolve such questions as whether or not prenasals are single segments. Although segments are clearly envisaged as being a part of AP, this has not been expressed in any formal way, so that gestures are said to form segments without there being any explicit indication as to precisely how this is done. For /x/, there are simply no explicit relationships between the component gestures comparable to those in (2) and hence no categorical processes which may act upon it.

Looking beyond FG and AP, different theories naturally differ as to both the type of basic units which should be employed, and the types of phonological relationships which should exist between them, making use of a wide variety of hierarchical relationships such as dependency and government between the basic units in their formation of segments. What is shared, however, is the recognition of the need for such relationships as illustrated in (2) to describe the categorical structures characteristic of phonology, and it is the lack of such relationships between gestures which Padgett claims leaves AP as not a theory of phonology but as solely a theory of phonetic implementation. This is a real problem, as although AP's ability to describe gradient phenomena has been seized upon with some gusto as a powerful and explanatory tool (e.g. Steriade 1991; Kohler 1992), its apparent lack of success at incorporating categorical information has led to a great deal of criticism (e.g. Clements 1992). In the following sections I will address some of these criticisms, assess the extent to which they are justified, and attempt to find some solutions.

2.2
The Categorical Content of AP

Given that each gesture in AP has a temporal as well as a physical dimension, we should be able to refer to the internal timing of gestures in order to specify invariant patterns of coordination between them. The mechanisms for such coordination have been made more explicit in Browman & Goldstein (1988). Here the number of points in each gesture available for intergestural coordination is, perhaps arbitrarily, restricted to three - onset, target and offset. These refer respectively to the starting point of the gesture, the period during which the goal of the gesture is considered to have been achieved (which need not be the same as the maximum degree of constriction), and the point at which movement out of the gesture takes place. Using these we can phase one point of gesture A with another point of gesture B and by simple empirical observation determine which of the various possibilities are actually attested.

However, these coordinations have been assumed to apply between constellations of gestures rather than between individual gestures. Thus Browman & Goldstein discuss the coordination of /p/ with /i/ in the phrase piece plots, assuming that the various components of both the vowel and the consonant are already coordinated so that they behave as if they were single events i.e. segments, rather than coordinations of many events. For example, while e.g. /p/ consists of a bilabial closure gesture, it also consists of a glottal opening gesture, yet Browman & Goldstein provide no means of determining the coordination presumably present between these two separate gestures, and although it might seem adequate to simply state that the LIPS and GLO gestures for /p/ completely overlap each other, we in fact require a far more rigorous definition of the possible patterns available, and why it is that this particular pattern is chosen here. More complex segments, such as prenasalised stops, affricates, labial-velar stops, or clicks, would inevitably have still more complex patterns of intergestural coordination at the segmental, categorical level, and it is hard to see how the simple coordinative patterns of e.g. Browman & Goldstein (1991) could be extended to handle cases such as these.

This is in essence the criticism highlighted by Padgett (1991) and discussed above. The problems lie in two areas. Firstly, we must determine what type of phonological relationships hold between the gestures that make up a single segment such as /p/, and differentiate them from various other types of relationships which may exist, such as those between onset head and dependent (e.g. /p/ and /r/ in 'pry'), or between a coda and a following onset; although the two types of relations might very well be capable of being expressed in the same terms, their functions are quite different. Secondly, we must identify the coordinative consequences of these phonological relationships. As Clements (1992) points out, unless the possible types of coordination are constrained in some systematic fashion, the number of potential contrasts expands exponentially with each gesture added. For example, given just two gestures to coordinate, we can phase the onset of gesture 1 with the onset, target or offset of gesture 2, and the same possibilities extend to the coordination of gesture 1's target and offset, giving us as many as nine possible outcomes. In addition, a fact not noted by Clements, there seems nothing in principle to prevent us from phasing e.g. gesture 1's onset to gesture 2's onset, and in addition phasing gesture 1's target to gesture 2's offset. This would possibly result in the duration of gesture 1 being increased to a certain extent, and at the very least would result in its extending beyond the period during which gesture 2 is active. If this were desirable we would want to ask to what end it could be made use of, and if not, how we might prevent it.

Browman & Goldstein's reply to this criticism is to claim that these differences in phasing patterns may in fact be necessary in order to adequately describe the data, citing Steriade (1991) as support.[3] While this may be true to a degree, it seems difficult to believe that two gestures could be distinctively coordinated in so many ways, and the mere possibility of so many patterns being distinctive is fragile ground on which to build a theory. The problem is compounded by Kingston & Cohen (1992), who make the point that vastly fewer (distinctive) phasing patterns are in actuality attested, and that FG seems to allow all and only these attested patterns, making it a true theory rather than merely a description, while AP fails to even adequately describe these patterns.

Kingston & Cohen raise in particular the problem of the representation of aspiration, citing the example of Icelandic as a language which can only be

[3]Steriade's proposals will be examined in chapter 5

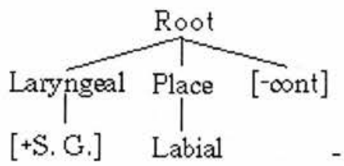
adequately described by a theory where ordering cannot be phonologically distinctive. They claim that AP's inherent duration of gestures would force it to be overly explicit in regard to the coordination of laryngeal and supralaryngeal gestures, so making it incapable of expressing the well-attested fact that while any one language may have both pre- and postaspiration, there is no language which phonologically contrasts the two. In (3) below I give data from Scottish Gaelic. We can see that although we find both types of aspiration, there are no words contrasted solely by pre- vs postaspiration. Word-initially, we find both voiceless unaspirated and voiceless postaspirated stops, but medially the contrast is between voiceless unaspirated and voiceless preaspirated stops. In other words, preaspiration precludes postaspiration and vice-versa.

(3)

a.			b.		
ball	[pauɫ]	n. 'ball'	palla	[p^{h}aɫa]	n. 'rock shelf'
dall	[tauɫ]	adj. 'blind'	toll	[t^{h}ouɫ]	n. 'hole'
gall	[kauɫ]	n. 'foreigner'	call	[k^{h}auɫ]	n. 'loss'
c.			d.		
cupan	[k^{h}u^{h}pan]	n. 'cup'	cab	[k^{h}ap]	n. 'open mouth'
bata	[pa^{h}tə]	n. 'stick'	badan	[patan]	n. 'tuft'
tacan	[t^{h}a^{h}kan]	n. 'short while'	daga	[takə]	n. 'pistol'

Kingston & Cohen claim that this situation is precisely what FG predicts (a partial representation of /p^{h}/ is given in (4)). Again we find that any gradient information is excluded, only categorical information being considered relevant for a phonological description. FG assumes that all features in the tree are considered to be phonologically unordered, and in general produced simultaneously, no one feature to be offset with respect to another. The fact that the laryngeal features extend beyond the supralaryngeal features is taken to be phonologically unimportant, on the premise that if no phonological rule refers to such ordering, then no such ordering should be present in the phonological representation. What is it then that would allow us to derive the correct phonetic ordering from such an unordered phonology?

(4)



Kim (1970) proposed that what distinguished aspirated from non-aspirated stops was the size of the glottal opening rather than voice onset time (VOT). Measuring his own (Korean) speech, Kim noted that while the glottis was opened wide in the production of lax (voiceless unaspirated) stops, the opening was far wider for the aspirated stops. From this he surmised that it was the difference in the size of glottal opening that was the important factor. As the glottal opening was wider for the aspirated stops than for the unaspirated stops, the amount of time taken for the distinctive voicing of the following vowel to be achieved would be that much greater, giving the impression of aspiration. This would allow aspiration to be modelled without the need for any complex timing relationships between the glottis and the supralaryngeal articulation.

The use of [+spread glottis] in the representation of $/p^h/$ in (4) is a reflection of Kim's position. As the glottis and lips are not on the same level in FG, there can be no phonological ordering, but Kim's analysis obviates the need for this. Such an analysis is clearly a very welcome one, having the added advantage that we do not have to refer to some arbitrary feature such as [+aspiration]. However, the use of [+spread glottis] brings with it the problem that the phonetic reality does not always match the phonological description, most particularly in that the glottis may be 'spread' without giving rise to aspiration. While the glottal opening for aspirated stops is indeed large, the opening for voiceless fricatives is often appreciably larger without giving rise to aspiration (Catford 1977). The reason for this, of course, is that while fricatives need a wide glottis to allow sufficient airflow for the frication in the oral cavity to be audible, they are simply not intended to be aspirated. The wide glottal opening is there purely to assist the frication. Voiced fricatives indeed often have very little noticeable voicing, and may be produced with a completely open glottis, which may be as wide as the opening for aspirates (Ball 1984). As Catford notes, the better voiced a fricative, the poorer a fricative it becomes, as the voicing restricts the amount of airflow into the oral cavity and thus the amount of airflow available for good frication.

This does not immediately rule out [+spread glottis] as the feature responsible for aspiration, it merely means that we cannot automatically equate a wide open glottis with aspiration but must take into account the overall environment in which it is found. Given a contrast between e.g. /t/ and /t^h/, we might wish to claim that the contrast is one of a relatively more open glottis for the aspirated /t^h/ compared to a relatively less open glottis for the unaspirated /t/; [+spread glottis] would then not necessarily refer to a specific degree of opening - no threshold would need to be crossed in order for us to say that a segment was unequivocally [+spread glottis] - but would be comparable to the use of [+voice] for the voiced stops of English which are often at best only partially voiced, but which contrast quite clearly with the voiceless series where voicelessness is complete. However, this would seem to be no better than the use of [+aspiration], especially as the same relatively more and less open glottal opening could be found as a contrast for fricatives such as /s/ and /z/.

These facts might be explained away by reference to the fundamental contrast in the nature of stops and fricatives, and Kingston (1991) has gone some way along this line. However, the use of [+spread glottis] seems unreasonable when faced with languages which contrast aspirated and unaspirated segments but which clearly do not rely on differences in the degree of glottal opening. For example, Shuken (1980) shows for Scottish Gaelic that unaspirated stops can have glottal opening gestures as wide as, if not wider than, aspirated stops. Similarly, Ladefoged & Maddieson (1996) note that voiceless unaspirated and voiceless aspirated stops in Icelandic may also show no difference in the size of glottal opening, and they suggest that the crucial factor determining presence or absence of aspiration is instead the timing of glottal opening with respect to the oral closure gesture. If timing rather than the size of glottal opening were to prove to be primary, FG would be left with a representation which fails to capture the facts correctly. Rather than using the size of glottal opening, i.e. [+spread glottis], what we require is the incorporation of ordering. Of course, FG cannot incorporate such information without drastically reassessing its theoretical basis, while AP not only can but must.

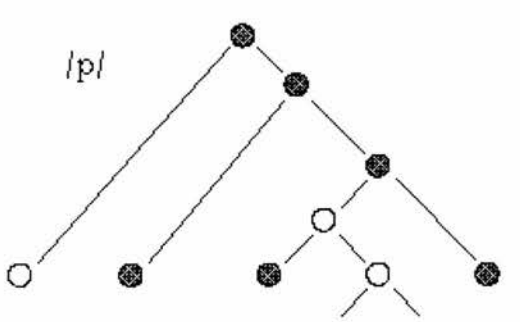
There are also theory-internal problems to consider. The use of [+spread glottis] is the only feature of its type in FG. If timing is the essential component of aspiration, then [+spread glottis] is the only feature to involve intrinsic ordering. Further, how are we to use it to represent voiced (modal or breathy) aspirated stops, such as those found in many Indo-Aryan languages, as [+spread glottis] seems to carry [-voice] with it? The answer to this is to employ a wider range of laryngeal features for languages which make use of [+spread glottis], so that languages such as Icelandic, which make a two-way contrast between voiceless aspirated and voiceless unaspirated stops, have far more complex laryngeal features than languages with a two-way voiced versus voiceless contrast. The question then arises as to why noe type of contrast should be represented as very much more complex than the other.

In conclusion, the use of [+spread glottis] for the feature of aspiration in FG is inadequate from both a phonetic and a phonological point of view. This does not in itself solve the problems facing AP, as the criticisms regarding the timing relationships between gestures are still valid, but it does show that FG's claim to more adequately describe the data does not necessarily hold good. Both theories fail, albeit on different grounds. However, the advantage which AP possesses is that it is able to capture the timing aspects of aspiration in a quite straightforward manner. All that is needed is to constrain the relationships in a principled way so as to generate all and only those patterns actually attested, and in addition to exclude patterns which we do not find. The balancing act is to allow certain relationships to reflect what is usually thought of as categorical information, while not outlawing other relationships which may be needed in order to capture the more fine-grained 'phonetic' information, and at the same time keeping the two domains distinct. In other words, we have to use the same primitives to create the phonology as well as the phonetics, and I will attempt to do this in this and the following chapters. To begin with, I shall examine more closely the relationship between the laryngeal and the supralaryngeal articulations, as expressed in AP, and will attempt to challenge Kingston & Cohen's criticism.

2.3
Creating Categorical Structures

I give in (5) below the AP representation of /p/ in terms of the Vocal Tract Hierarchy. The relationship between AP and FG should be readily apparent from a brief comparison between (5) and (4). The chief elements of each, if examined theory-externally, are roughly the same (though this may reflect the extensive common ground between phonological theories in general rather than any intentional similarity). For a segment such as /p/ there are basically three prime characteristics, namely the labial location, the complete constriction of all airflow, and the open glottis. This is, of course, encapsulated in the common description of /p/ as a voiceless bilabial plosive; to change any one of these descriptors would be to describe another segment. For example, if we changed the value of GLO to **cri** we would have /b/; changing the CL to Tongue Tip would give us /t/. Regardless of theoretical standpoint, any theory will perforce need to describe the same events.

(5)



While many of the same characteristics are thus incorporated into all phonological theories, the prominence given to each property differs from theory to theory. For example, since Lass (1976) more prominence has in general been given to the division between the laryngeal and supralaryngeal,[4] both being treated as relatively independent systems bound together at a higher point in the hierarchy. This allows for processes to affect one of the levels independently of the other, and is reflected in both FG and AP where a process such as debuccalisation is simply stated given a theory with more than one level. In the change from /s/ to /h/ in many varieties of Spanish (Harris 1969), for example, we can simply delete the supralaryngeal articulation, leaving behind only the laryngeal, i.e. /h/. Without such a hierarchical structure such processes become random events.

However, processes such as debuccalisation obviously require a degree of categoricalness in the representation. In FG, the [laryngeal] node is linked directly to the root node, as is the [place] node, and deletion of [place] will not affect [laryngeal] in any way as we have already seen (see (2) above). In order to describe such a process in AP we must also provide explicit phonological relations linking the laryngeal and supralaryngeal gestures in a discrete categorical way, whilst allowing the link between them to be broken, or simply not present. This in turn means that we need some means to describe this link. This leads us to ask what is it about the representation of (5) that allows us to derive the necessary information, and what precisely this information is.

For stops, as the name implies, the characteristic event is complete closure. For /p/ the lips come tightly together to block all airflow, resulting in a build-up of pressure behind the closure. This period of silence is what links all stops - in featural terms they are non-continuant, non-sonorant - though it is not all that links them. In word-initial position, the build-up of pressure results in a highly salient burst when the closure is released, and it is primarily this burst which allows the hearer to

[4]The question of whether we should in fact have a supralaryngeal node and not just a place node (e.g. McCarthy 1988) is not one which I shall address directly here, as the representation of structure in AP does not suffer from McCarthy's criticisms (though see the discussion of Bird (1991) below).

correctly identify the stop. This release, though, is not strictly necessary. In many languages, including English, stops in coda position, or word-finally, need not be released. Their identity is signalled by the formant transitions from the previous vowel, which, although not so salient as the release burst, are enough to ensure accurate transmission and reception.[5] Given that the release burst is not always necessary, we are left with only the period of silence or closure as the defining characteristic of stops.[6]

This closure of course is not unique to stops. Laterals and nasals, amongst others, also share the same closure, so that /t/, /l/ and /n/ all have Tongue Tip closure of roughly the same degree and location, and can act as a natural class as in e.g. Basque (Hualde 1991). It is the presence of nothing but closure that places /t/ in a natural class apart from /l/ and /n/. This is reflected in the gestural scores in (6), where both /l/ and /n/ include the oral portion of /t/ as part of their representations. What we do not see is how the TT closures on the one hand, and the VEL and TB gestures on the other, are phonologically related to each other. Indeed, we have no means of knowing whether any phonological relationships exist between the gestures at all, as it is assumed rather than made explicit that somehow we have categoricalness, and it is only the presence versus absence of gestures that distinguishes the three segments (cf. Browman and Goldstein 1992a). There appears to be no way of knowing how the various relations between gestures differ, if in fact they do. For example, is the VEL : **open** gesture in (6b) dominated by the TT : **clo** gesture in the same way in which [laryngeal] dominates [voice], or vice versa, or neither? And whatever the phonological relationship between them, is it in any way different from that between TT : **clo** and GLO : **open** for /t/?

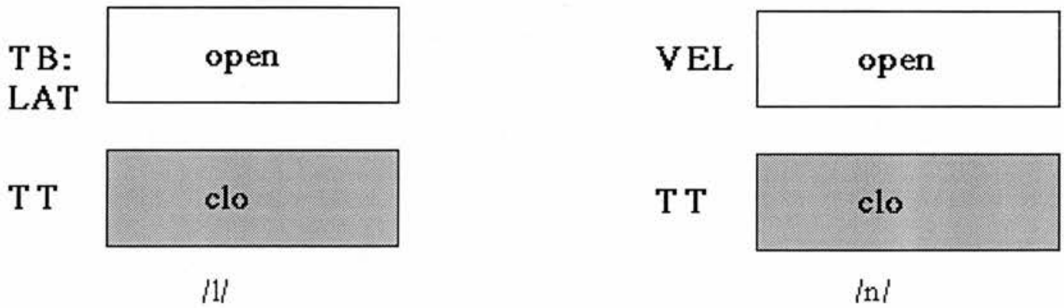
---

[5] In fact, the weakness of these formant transmissions, and the weakness of coda consonants in general, can lead to neutralisation in this position, e.g. Malay (Durand 1987).

[6] Release has been incorporated into a number of theories (e.g. Steriade 1992, Harris 1991). For example, Harris suggests that the /p/ in 'kept' should be phonologically distinguished from the /p/ in 'pep', suggesting they show differences in phonological behaviour.

(6)

(a) (b)

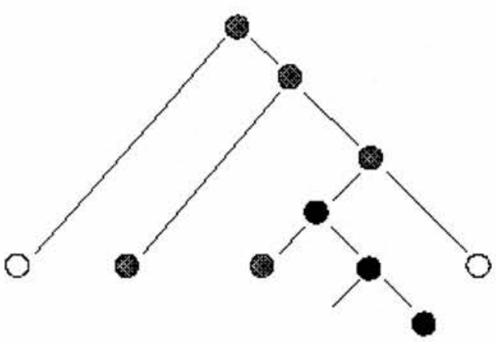


Recapping, in FG and other theories, there are mechanisms which tell us that the various components of /l/, /n/, /p/ etc. behave as if they formed a single unit, i.e. a segment, but there is nothing in the representations of (5) and (6) which provides us with the same kind of information. At best, we can perhaps account for instances where e.g. GLO : **open** by itself forms the segment /h/, as there are no other gestures with which the glottal gesture must combine and thus no problem in specifying intergestural relationships. However, this would leave us able to account for only a highly impoverished inventory of segment types. Returning to stops, if we cannot refer to categoricalness, then silence rather than closure is all we are left with, i.e. the acoustic effect of the gesture rather than the gesture itself, which seems a little at odds with the proclaimed articulatory nature of the phonology.

However, given the existence of affricates, we cannot claim even silence as exclusive to stops. We are left with an impossible position, for if we cannot provide a proper definition of categoricalness, we cannot truly say what exclusively distinguishes a stop from a lateral or a nasal, or from any other 'segment'. We can point to representations such as those in (6) and show how presence versus absence of gestures reflects some intuitive notion of what a segment is, but if we cannot show how TT and VEL are linked so as to form all and only /n/ we can do no better than to say that here we have two gestures linked in an arbitrary manner which happen to occupy roughly the same physical and temporal space and in so doing happen to constitute a regular pattern. Without being explicit about the nature of the relationship it would be no more arbitrary to classify the same gestures in the French 'tant' [tã] as a segment, even though the VEL gesture is not linked in any meaningful way with the TT gesture. In other words, no explicit formal link is no better than no link at all.

Although we know that we can represent /t/ in terms of a TT : **clo** gesture, we are left with little to say about its relationships to the gestures around it. The situation is not improved by referring instead to the Vocal Tract Hierarchy rather than the gestural score. (7) shows the same problems that we saw above for /p/. In particular, the presence of preaspiration or postaspiration could not be indicated in such a diagram, as all we can say is that we have an open glottis and a TT closure gesture, not how these two phase with each other. This is a result of the VTH being a snapshot of the state of the overall vocal tract at any one instant, rather than a phonological object in itself.

(7)

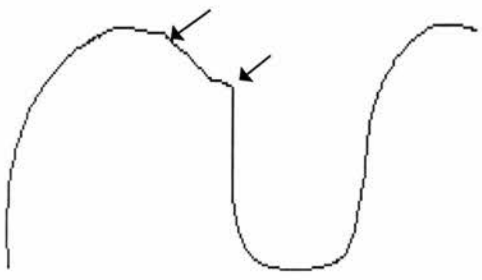


There are some further problems with diagrams of the type in (7), connected with their claim to be snapshots of the overall physical state of the vocal tract at any one point in time. Boyce et al. (1990) refer to the varying degrees of velum height in non-nasal consonants, in connection with earlier chain models of phonetics. They note that the velum is typically most tightly raised for plosives, less tightly so for fricatives, and progressively less so until with some speakers the velum can be more or less fully lowered for vowels. The reason for this is straightforward enough, in that stops with leaking airflow through an open velum would make poor stops - but good nasals - while the velum can be a little lower for fricatives without causing frication to cease. Boyce et al. go on to deal with two phrases which lend support to the AP approach to phonology, but which also raise questions for it.

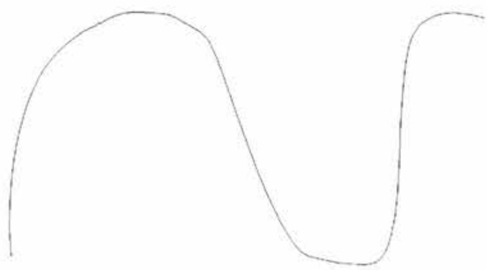
In (8a) (adapted from Boyce et al. 1990) we see a representation for velum height during the word **lansal** in the phrase it's a lansal again, and in (8b) we have another representation for the same word in the same phrase, though with a clearly different pattern. In (8a) we see that we have two separate periods of velum lowering (the initial points of the lowering movement are indicated by arrows), the second and larger one being due to the /n/ in lansal, but Boyce et al. attribute the first and smaller lowering movement to the /s/ in it's a, though non-contrastively. In (8b) however we see apparently only one lowering movement, but we can also see that it is a larger movement and Boyce et al. assume that what we have is in fact two overlapping lowering gestures which appear as one larger gesture (cf. Browman & Goldstein 1986). This of course agrees with the AP approach.

(8)

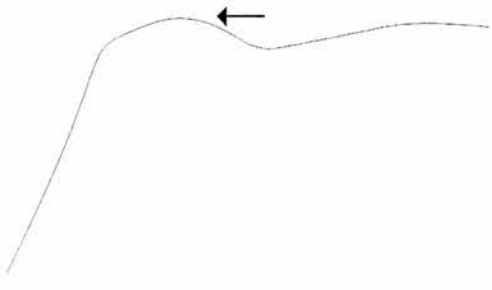
(a) It's a lansal again



(b) It's a lansal again



(c) It's a lasal again



In (8c) we have the representation for the velum height in the phrase it's a lasal again, i.e. the same phrase but without the nasal. It could be argued that the lowering gesture in (8a) might have been due to some anticipatory nasalisation of the following nasal, but (8c) shows quite clearly that even without this nasal we still find a lowering gesture (onset of lowering again indicated by arrow). This gesture is obviously connected with the /s/ of it's a as noted above, but how are we to describe it and its relation to the apparently active TT gesture?

AP describes controlled and planned articulatory movements within the vocal tract, and it is these that the gestural score and the VTH are designed to reflect. A velum lowering gesture is such a planned gesture, as in /n/ - it is actively controlled by a task dynamic equation which directs a set of tract variables towards specific goals, goals which are in part defined by their not being the neutral settings discussed in Chapter 1. Yet here we have a velum lowering gesture which presumably is not actively controlled by an equation, but which is instead controlled by another gesture, or by some separate set of language-specific values (Browman & Goldstein 1989). The intuitive basis for distinguishing the VEL gestures in /n/ and /s/ is clear, and they would generally be separated into the phonology and phonetics respectively, but in a partial-phonetic system such as AP we cannot simply draw an imaginary line between the two domains, but must provide a mechanism for distinguishing the two while at the same time recognising that both are of the same type. Again the problem seems to be the reverse of those facing FG, in that we have the ability to make fine grained phonetic distinctions, some of which are highly relevant to the phonology, but we have no way of distinguishing between what is categorical and what is not.

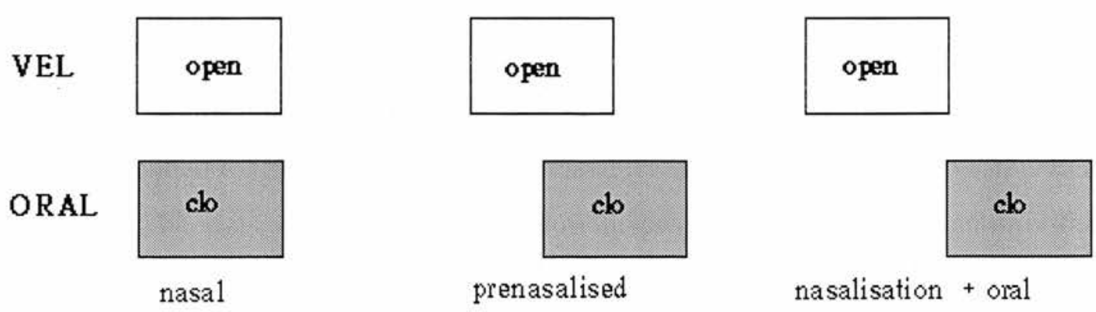
Mattingly (1990) makes similar comments when he discusses the value of introducing full specification into AP. Browman & Goldstein themselves claim that AP is 'inherently underspecified' (Browman & Goldstein 1989), but Mattingly cites evidence such as non-contrastive velum lowering to suggest that in a more comprehensive system each and every articulatory movement would be included. We could then encode such phenomena as velum lowering for /s/ indirectly. Primarily we would have a TT : **cri** gesture which would be under the direct control of a task dynamic equation, but secondary to this would be a velum lowering gesture which would be dependent on the primary critical gesture and controlled by it. The velum lowering gesture, while not being a neutral one, is still distinguished from the 'primary' active gestures by means of some as yet unspecified relationship between the two.

Again we note the familiar problems arising, those of categoricalness and gradience. Without the categoricalness we can not really claim that mere presence or absence of gestures can give rise to contrasts, rather we can make the lesser claim that it could *possibly* give rise to contrasts. Until or unless we can formally show what the relationships between gestures are it is useless to claim that presence or absence of these relationships can affect anything. Until we can define whether all non-neutral gestures are active or only some, it is useless to claim that active gestures have any special relevance. In other words, without the basic skeletal frame of a series of explicit and constrained coordinative patterns, the process of intergestural phasing becomes arbitrary and meaningless, and AP will remain unable to do anything other than describe what it sees without explaining what is going on in the speaker's mind (Ladefoged 1991).

What possibilities are there for increasing the explicitness of the gestural phasing? We can see that without the introduction of mechanisms for producing categoricalness there is indeed no means of rebutting the criticism of Kingston & Cohen (1992) that AP is unable to rule out the possibility of pre- and post-aspiration in positions of contrast in any one language. The fact that the FG analysis is itself open to criticism is of no real relevance, as the relative success or failure of a theory should be measured in terms of how well it describes the data and not by how poorly some other theory performs the same task. However, while it may appear from the above discussion that categoricalness as such has not been of prime concern to Browman and Goldstein, this should not be taken to mean that it has been entirely absent from the theory. In fact there has been a great deal said about the putative categorical nature of AP, but I shall show that the consequences of these claims are not ones we should want.

(9) below (adapted from Browman & Goldstein 1991) shows the three possible degrees of overlap which are claimed to be phonologically relevant in AP. In (9a) we see the results of complete overlap, giving us the 'segment' /n/. Here there are two active gestures, namely VEL : **open** and TT : **clo**. The actual value of voicing in speech can, of course, take many forms, from modal voicing through whisper to full voicelessness, but every supralaryngeal gesture will overlay glottal activity in some way. AP makes the claim that when an articulator is not actively involved in performing a gesture it takes up (or attempts to take up) a neutral rest position, although this neutral target may not actually be achieved. As most segments, including vowels, are voiced, and the neutral setting for language as opposed to breathing seems to involve a position suitable for modal voicing (Laver 1994), AP assumes that voicing is the neutral setting for the glottis. As the gestural score is intended to show only those gestures which are active, the assumption is that /n/ will only need its active gestures shown in the gestural score, and we can infer the presence of voicing from our knowledge of the neutral settings. However, this again raises the question of what is active and what is neutral, like the non-phonological differences in velum height noted above. It may very well be that we should be able to refer to such non-distinctive voicing in phonological rules (Lombardi 1990) yet the gestural score suggests otherwise.

(9) **(a)** **(b)** **(c)**



However, this brings us back to the criticisms of Mattingly (1990). Given that we have two types of non-predictable information - the gestures themselves and the neutral settings dependent on the gestures, neither of which are strictly physically predictable - we have two main choices as to how we might distinguish between them. The first choice is to directly encode the dependence of the neutral settings into the theory, as discussed above, by allowing active gestures to control the various settings. The second is to assume that the two types belong to different domains, and thus include only one of them in the relevant part of the model; in other words we ignore part of the information, or underspecify. It is the second option which Browman and Goldstein appear to choose, but again I would claim that in a fuller and more elaborated theory it is the first option which should be favoured. There seems to be a distinction between neutral settings which are true for speech as a whole, and other neutral settings which are dependent upon the presence of specific active gestures, yet with full specification this distinction could be eliminated.

Thus the choice of which information to present in the gestural score remains to a large extent arbitrary until we can show how the non-neutral 'gestures' are different from other non-neutral and non-predictable information. Secondly, the pattern of overlap of the VEL and TT gestures for /n/ also appears to be arbitrary. The various patterns seen in (9) form a coherent system in many ways. /n/, as already noted, consists of two completely overlapping gestures, forming what most observers would regard as a single segment. We also have two other forms of overlap possible, namely partial and minimal, which we see in (9b) and (9c) respectively, only the latter of which would generally be regarded as a segment. (9b) shows /n͜d/, a prenasalised voiced stop, while (9c) shows a sequence of a nasal vowel and an oral stop. Immediately we see that these are two very different kinds of objects, and we would not normally consider (9c) to form a segment or any other coherent unit at all. What connection, if any, is there between these forms which allows or compels us to represent them in the same terms?

Overlap is the most pervasive of all relationships in AP, in fact any relationship between two gestures necessarily involves some overlap between them. Browman & Goldstein, as noted earlier, use this pervasiveness to establish some invariance in overlap between vowel and consonant gestures to reflect such positional relationships as Onset-Nucleus, or Nucleus-Coda. In the phrase piece plots both initial /p/'s would generally be considered to be in the onset and to have basically the same type of relationships as each other with their respective nuclei, but at the same time we would never consider e.g. /p/ and /i/ in piece to constitute together a segment of any kind. As discussed above, it is the various gestures which go to make up /p/ which constitute segmentality, and similarly /i/ is a categorical constellation of gestures which act as if they were a single unit, a segment. However, given that we use the same processes of overlap for each of (9) we make the implicit assumption that (9c) is of the same basic type as both (9a) and (9b). By this the notion of a segment becomes irrelevant. Is this a useful assumption?

The various distinct types of overlap identified by Browman & Goldstein are comparable to the quantal nature of the various degrees of constriction degrees (cf. Stevens 1972). Let us assume that we have two overlapping gestures as in (9). Presumably, we begin with complete overlap, which will give us a particular complex set of sounds which we can refer to as a single segment, and then we can shift the gestures with respect to each other. Eventually the two gestures will shift far enough apart so that they cross some perceptual barrier which a listener would perceive as representing some new type of segment. This we interpret as a shift from complete to partial overlap.[7] However, there is nothing to prevent us from sliding the two gestures still further apart until they show only minimal overlap, and at this point another perceptual barrier would be reached. Two gestures showing minimal overlap are interpreted not as a new type of segment, but as a sequence of two segments. Despite this, we would expect that the two gestures would still behave as a unit of a kind, though it would be a different kind of unit from that found with the other two types of overlap. Finally, we could slide the gestures apart still further until they no longer overlapped at all, and then presumably the listener would perceive no unit of any type whatsoever.

[7] As noted earlier, AP might also wish to make use of the various stages of overlap which lie between complete and partial (Steriade 1991), but the capability to make formal distinctions between the 'end' stages and those between the ends is not yet present within the theory.

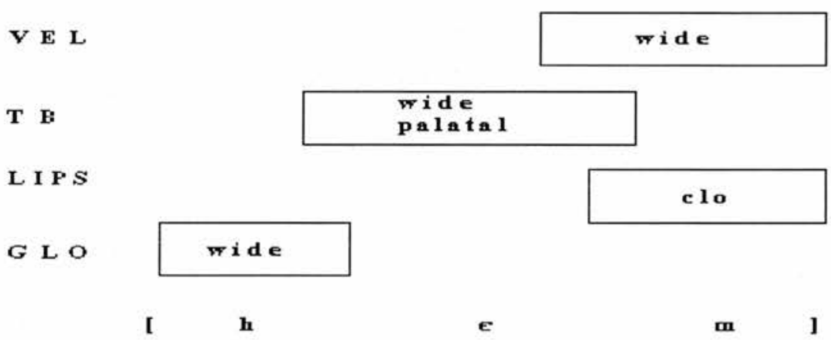
The question which arises is why is a barrier crossed as we move from partial to minimal overlap? There is nothing in these two relationships which inherently suggests that such a perceptual leap should take place. In fact, the sole basis for such a distinction seems to be based on the perception, which then feeds the interpretation of the relationships. In other words, it is not strictly the intergestural relationships which are important, rather it is the perceptual patterns which they allegedly give rise to which are of any linguistic significance, resulting in distinctions based more on acoustics than on articulation.

Again, the problem lies with AP's failure to distinguish adequately between categorical and gradient relationships. The gestural score represents distinct relationships in terms of the same structures, but frequently blurs these distinctions, or at least does not enable their different statuses to be determined solely by reference to the gestural score representation. For example, if we look at the example of <u>perfect memory</u> given in Chapter 1, we assume that we have categorical relationships defining all the consonants as consonants and all the vowels as vowels. At the same time, the flexibility of the gestural approach allows us to slide gestures with respect to each other and give the appearance of the creation of new categorical structures where none such in fact exist. Thus, the feature deletion accounts traditionally employed to explain such forms as <u>perfec memory</u> are shown to be inadequate to explain the data, while AP deals with these phenomena with ease. The same principles apply to assimilation such as in <u>seven plus</u> → <u>sevem plus</u>, where the coronal gesture for the /n/ of 'ten' is still present in the assimilated form, showing that we have a gradient phenomenon which has the external appearance of a categorical phenomenon.

But do AP and the gestural score really give us an adequate representation? The claim that such assimilation is gradient implies that the 'segments' formed by such processes are themselves gradient rather than categorical structures. Byrd (1992) implies as much in her discussion of the degree of overlap necessary before assimilation is perceived. Complete overlap is by no means needed, so assimilation may be perceived at perhaps 60% overlap, but this leaves us with a great deal more possible overlap, up to and beyond complete overlap.[8] Presumably we would want to distinguish between such cases of /m/ in sevem plus and a 'normal' /m/ as in hem, but the lack of any categorical information in the gestural score prevents us from doing so.[9] (10) below shows partial representations of these two forms. It is clear that the presumably categorical status of the VEL and LIPS gestures for the /m/ in (10b) appears to be identical with that of the same gestures in the assimilated form in (10a). In itself this is no bad thing, as it explains why it is that assimilation is perceived, i.e. why the listener cannot tell the two forms apart - they appear identical because on some level they are. However, while AP is making the claim that what might appear to be a categorical change is in fact a gradient one, it is unable to show this in the Gestural Score.

(10)

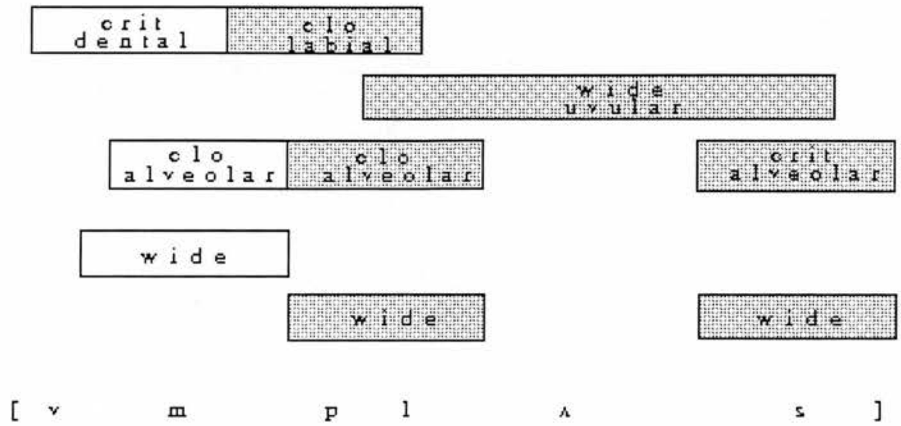
(a)



[8] Though it may sound paradoxical to claim we can go beyond complete overlap, which we can think of as being 100%, overlap of e.g. 120% would simply imply that what was once C^2 in a $C^1\#C^2$ sequence now not only completely covers C^1 but begins before it.

[9] The alternative view is to claim that there is in fact no difference between the forms. In that case, we would be entitled to ask why it is if 60% overlap is enough to signal assimilation and thus segmentality, AP models 'normal' segmentality as 100% overlap. In addition, although Byrd uses different phasing relationships between consonantal gestures to model assimilation, it seems equally plausible that in cases such as sevem plus the LIPS gesture is in fact phased earlier with relation to the TB gesture for the preceding vowel. If this were so, the representation in the GS would be still more inadequate.

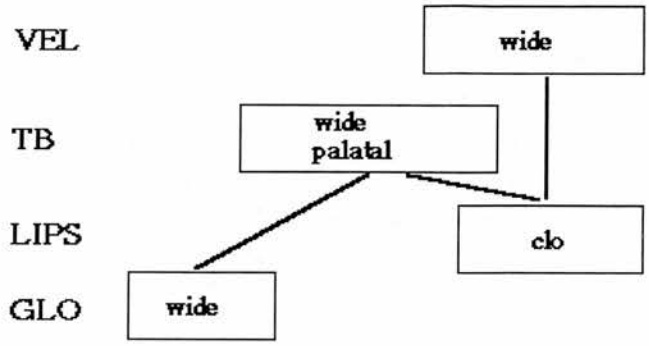
(b)



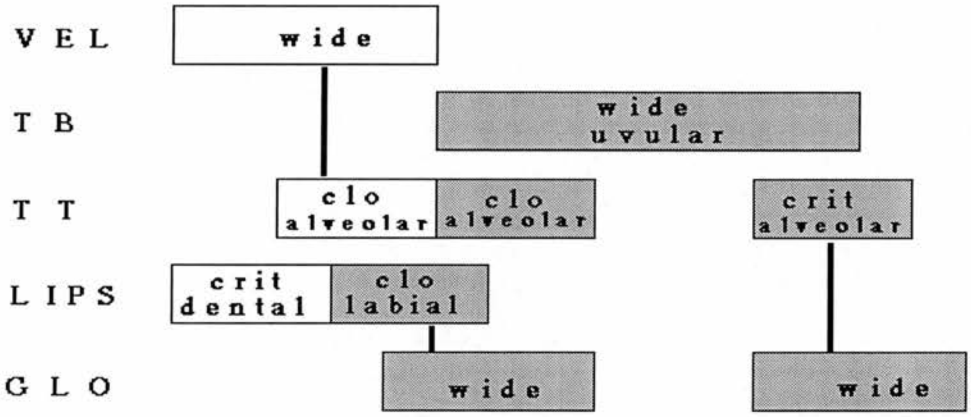
An alternative to the straight gestural score is the point notation employed in Browman & Goldstein (1989). In (11a,b) I show the forms in (10a,b) in an adapted form of point notation. The lines connecting the various gestures suggest some as yet unspecified type of categorical information, perhaps comparable with the association lines of FG. The straight vertical lines are intended to represent segmental information, while the diagonal lines encode intersegmental information such as Onset - Nucleus. The important point is that for the assimilated forms we can omit any connecting lines, and interpret such absence of connections between overlapping gestures as implying gradient rather than categorical information (the phasing between the consonantal gestures and /ʌ/ of 'plus' is not shown). Hence (11b) has what we might refer to as a 'real' /m/, while (11a) has a 'false' /m/ derived by assimilation, i.e. sliding of gestures. Unfortunately, without a real explanation as to why some relationships are categorical while others are not, and what relationships are specified by the connecting lines, we are left with only a diagram on a page, not an explanation. The earlier criticisms still apply, and we have yet another area in which the lack of categoricalness leaves us unable to provide a convincing analysis.

(11)

(a)



(b)



What, then, are we left with? Without a method of distinguishing between categorical and gradient information there appears to be no way to exclusively refer to segmentation of any kind. There is a clear need to be able to refer to some clusters of gestures as behaving as if they were single units. Browman and Goldstein do assume that such a position already exists within AP, but it is clear that such a system is only apparently present. Without categoricalness it is of little worth to show how certain phenomena are distinguished by being gradient rather than categorical, as we would have no categorical point of reference from which to distinguish them. To return to (9) the difference between (9a,b) and (9c) is that the gestures in the former form categorical relationships of a kind clearly different to the relationships formed in the latter. How these differences are to be encoded is another matter.

In the following section I will take a closer look at the Vocal Tract Hierarchy, the main mechanism in AP for encoding segmentation, showing how it too fails to provide a basis for distinguishing between gradient and categorical structures.

2.4
The Representation of Complex and Contour Segments

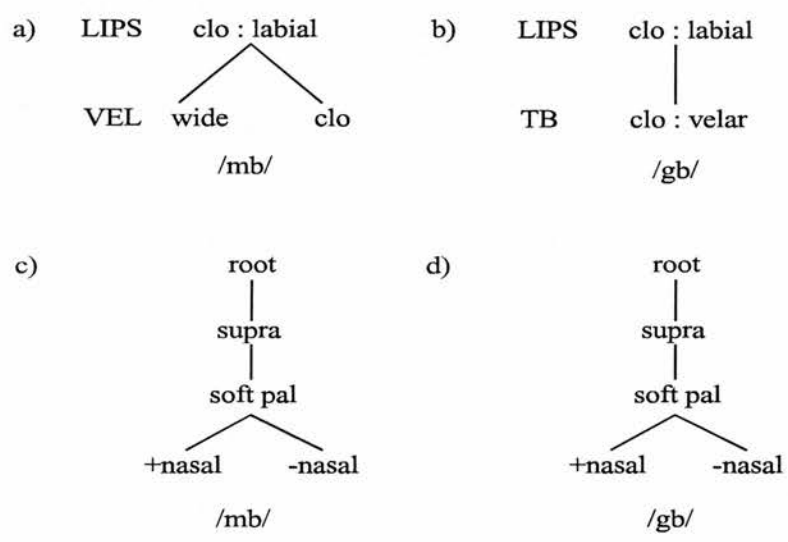
The range of segment types found in natural languages is bewildering to say the least, encompassing a vast range of sounds, only a few of which are to be found in any one language. One of the things which make the study of these sounds so appealing is that while it is clear that there are rules governing their behaviour, it by no means clear what these rules are - hence the proliferation of theories battling to cope with these facts. Each theory will claim to provide a more elegant and descriptive analysis than the others, but of course what we usually find is that this elegance often does not extend beyond the relatively small set of data which originally gave rise to the theory. This does not, and should not, mean that we should not attempt to extend the range of our theories, as it is only by doing so that we can explore their limitations and at the same time provide new insights as to the nature of the data. The shortcomings of one theory in turn leads to the birth of another, incorporating the insights made by its predecessor, but at the same time hoping not to make the same mistakes. It would be an error to believe that AP is anything other than a natural part of this process. However, it seems a feasible goal to see how far the set of data which AP can handle can be extended, and at the same time to see which areas it is patently unsuited for, as such areas are bound to arise.

We have seen briefly above that there are particular areas of phonetic/phonological study which AP claims to represent better than other theories, but in doing so leaves itself apparently unable to handle areas which these same other theories appear to handle with ease. Many of these problems, though, can be reduced to problems of interpretation, and to the particular manner of representation chosen by Browman & Goldstein. As such it would profit us to look at other ways of incorporating AP's insights which avoid the shortcomings we have already noted, while maintaining AP's advantages over other theories. Many of these advantages of AP concern the explicit incorporation of timing into the representations, as has been noted a number of times already, but in fact it remains to be proven to what extent this is in fact an advantage over FG representations, especially in the light of more recent theoretical developments in FG.

Browman & Goldstein (1989) make strong claims as to the superiority of AP regarding the (graphical) representation of so-called contour and complex segments, and (12) below (taken from Browman & Goldstein 1989) illustrates some of the differences between the two theories. (12a,b) represent a prenasalised stop /mb/ and a complex stop /gb/ respectively in AP, and are clearly distinguishable from each other, both in the number and type of gestures involved and in the phasing relationships which hold between the gestures. Even at a casual glance (12a,b) would appear to represent two different segment types. Browman & Goldstein (1989) claim that in contrast, the FG forms in (12c,d) do not immediately appear to differ from each other - each shows a mother node dominating two identically branching daughter nodes - and that without further knowledge of the conventions of the theory we could not immediately class (12c,d) as representing different segment types. Assuming that these are in fact different segment types and that any theory should be able to distinguish the one from the other, AP, Browman & Goldstein claim, would appear to have a clear advantage over FG, but what is the justification for such a claim? To see if AP does indeed provide a more satisfactory description of the data it is necessary to examine what precisely is being represented.[10]

[10] The very existence of complex segments and contour segments as distinct categories has recently been called into question by a number of researchers in FG (e.g. Padgett 1991; van de Weijer 1993). Clearly, if this distinction were to be eliminated, it would leave AP in the position of being forced to distinguish segments which in fact should not be distinguished, or at the very least of being unable to show the fundamental identity of the two types. I shall return to this point later.

(12)



It was Sagey (1986) who provided the first real FG formulation of complex segments, citing a number of pieces of evidence for marking them as a separate, easily identifiable class. Simple segments, in FG terms, have only one articulator node under the place node, and any degree of closure specified for the segment automatically applies to that articulator. Complex segments, on the other hand, have at least two active articulator nodes, which do not necessarily share the same value for manner, for example, the labialised [gw] of Nupe, or the palatalised [tj] of Russian. Sagey discussed the problem of how to ensure that, for complex segments with conflicting manner features, the correct manner feature was percolated to the correct articulator node, a problem to which I shall return below. However, Sagey pointed to a number of characteristics of complex segments which seemed to demand that they be treated differently from contour segments.

Perhaps the most important difference between complex and contour segments is that the former (apparently) behave as if the two articulator nodes were phonologically coextensive, i.e. there is no phonological ordering between the two, while the latter behave as if the different articulators were both phonologically and phonetically ordered. This is seen most clearly in the relationships between the two segment types with other segments, as in coarticulation or directionality of spreading.

For example, the data in (13) below, taken from Piggot (1988), show how prenasalised segments can apparently behave both as nasals and non-nasals. Guaraní is there claimed to have both leftward and rightward nasal spreading which is blocked by (supralaryngeal) stops, and has completely nasal morphemes, completely oral morphemes, and morphemes with a nasal part followed by an oral part. As can be seen from (13), prenasalised stops block the spread of nasality, thus acting as if they were oral stops, and they also do not trigger rightward spreading. In other words, they act as if they were oral stops when viewed from the right hand side. This in turn implies that the ordering of the nasal and oral parts is a vital part of the segment, as any preceding or following segment must be sensitive as to whether a [+ nasal] or [-nasal] value is adjacent to it. Thus contour segments are distinguished by the presence of ordering at the phonological level.[11]

(13)

a. tupa	'bed'	e. tũpã	'god'
b. piri	'rush'	f. p̃ĩr̃ĩ	'to shiver'
c. haihu	' to love'	g. mãpẽ	'to see'
d. g^{w}ata	'to walk'	h. nũpã	'to beat'
		i. pãrã	'to embellish'

j. m͜baʔe	'to think'
k. hẽn͜du	'to hear'
l. mẽn͜da	'to marry'

Of course, without ordering we would be left with a segment with two simultaneous but mutually exclusive values, plus and minus [nasal], so in that sense we have no choice but to specify phonological ordering - if we did not, we would

[11]In fact Piggot (1992) presents an analysis of the same data without requiring contour segments as such (Piggot's proposals are discussed in chapter 3). Sagey (1986) presents data from Land Dayak as also showing evidence of ordering in contour segments, but these too are better analysed without such ordering. For the purposes of the present discussion I will assume that such ordering is present.

have a segment with contradictory and inconsistent values.[12] The situation with complex segments is somewhat different. With contour segments the relevant values are on the same branch of the geometry. For example, we have [+/- nasal] as above, or [+/- cont] for affricates. However, for complex segments we are dealing with different branches, albeit branches subordinate to the [place] node, so the problem of contradictory values being simultaneously present does not arise, as different branches are assumed always to be phonologically unordered. While the number of branching nodes is usually two, there is no evident phonological reason to prevent multiple branching of articulator nodes, as they could all be interpreted as simultaneous with no unwanted side-effects for the theory. The fact that such multiply articulated segments are extremely rare, if indeed they exist at all, is due to other factors, such as difficulties in maintaining perceptual distinctions (Ladefoged & Maddieson 1996).

The simultaneity of the articulator nodes is reflected in a number of ways. Sagey cites the Kru word for 'dog', which has the varying realisations [bwe] ~ [gbwe] ~ [gbe], i.e. labial-velar ~ velar-labial-velar ~ velar-labial. In other words, in gestural terms the exact phasing between the labial and velar gestures is irrelevant, as long as they overlap each other to some extent. This free variation in phonetic ordering is not the only evidence for lack of phonological ordering. While contour segments are reputedly characterised by their different effects on adjacent segments, depending on whether those segments precede or follow the contour, so complex segments are characterised by having *identical* effects on adjacent segments, no matter whether they come linearly before or after. This is most clearly seen in the form taken by prenasalised complex segments. In (14) below we see the effect prenasalisation has on the simple and complex segments of Margi (Hoffman 1963). Examples given are in Margi orthography, adapted to show both articulations for the complex nasals (e.g. Hoffman writes 'md' where I write 'mnbd'). For complex segments both major articulators spread to the [+nasal] node, resulting in segments with a doubly articulated nasal part followed by a doubly articulated oral part, strongly suggesting a

[12] Naturally this does not apply if [nasal] is a unary feature. Cf. van de Weijer (1993)

lack of phonological ordering. If there were ordering present we would expect either one or the other of the articulators to spread leftwards, but certainly not both.

(14)

Oral		Nasal	
a. bdàgù	'valley'	d. mnbdá	'to surpass'
b. bdzà	'foolish'	e. mndzànì	'to spoil'
c. ptía	'to reduce'	f. mnptsàkù	'to pick (up)'

In addition to this, Sagey quotes evidence from !Xóõ showing the same phenomenon from the right edge. Clicks in !Xóõ involve both dorsal and coronal closure, and as such satisfy the environments for rules which disallow front vowels after dorsal consonants, and at the same time satisfy the environment for a rule raising /a/ after coronal consonants. This of course suggests that clicks, which are complex segments, are simultaneously dorsal and coronal.

Looking back at (12) it is clear that branching in FG diagrams does not automatically imply phonological ordering. Rather it implies two opposite states - phonological ordering on the one hand, and phonological unordering on the other. The interpretation of the diagrams depends upon a knowledge of the kind of objects which can branch, in particular we need to know whether we have two different values of a single node, in which case we have phonological ordering, or two separate nodes which happen to share the same parent node, in which case we have no phonological ordering, as in (12c,d). This of course is not immediately apparent from the diagrams, where ordering and unordering are both represented by branching, a fact which forms the basis for Browman and Goldstein's criticisms. At the same time, this is a somewhat weak argument, given that contour and complex segments are formally distinct, and that this distinction is in fact indicated in the diagrammatic representation.

What, then, do the alternative AP forms in (12a,b) represent? With the central position of importance given in AP to the incorporation of time into phonological representations, any diagram which did not explicitly incorporate this information would be inadequate. However, both the AP forms given seem to suffer from a number of inherent problems which make them if anything less adequate than the FG forms in (12c,d) and fail to incorporate the timing dimension in any meaningful way.

To begin with, (12a) suggests that prenasalised stops such as /mp/ consist of three separate gestures, namely LIPS : **clo**, VEL : **open** (lowering the velum) and VEL : **clo** (closing the velum) and not two. There seems little to support such a claim, however, and much against it. Browman & Goldstein (1986) suggest that for instances of nasal spreading we would need a VEL gesture which consisted solely of lowering; this gesture would then spread leftwards, rightwards, or both, with the velum remaining lowered, presumably, until a VEL : **clo** ('blocking') gesture was activated. Without a closure gesture the velum would remain lowered, so the existence of a raising gesture is vital. This would reflect the fact that such cases of nasal spread seem to consist of a single nasal gesture, identical in all respects to the nasal gesture in e.g. /n/ except that it is longer (Benguerel et al. 1977). Given this, for prenasalised segments, there must necessarily be both a lowering gesture and a raising gesture, resulting in the form in (12a).

The identity of the blocking gesture, however, remains a mystery, as while supralaryngeal stops seem to always block nasal spread, they do not have any active velic closing gesture. In fact, the 'inherent underspecification' of AP suggests that non-nasal segments have no VEL gesture at all, which would result in nasal spread only being blocked by other 'normal' nasals. This is not the result we would wish for. This raises again the question of precisely what information should be included in AP's phonological representations and how different types of information should be encoded. In addition, it implies quite clearly that the oral labial portion of /mp/ would be phonologically different to a 'normal' /p/, as the former has an explicit

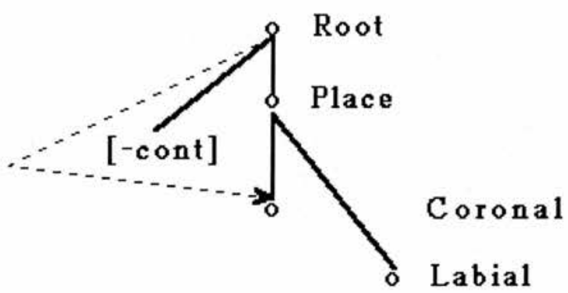
VEL : **clo** gesture associated to it while the latter does not. I know of no evidence to support such a distinction.

It seems that there is really no justification for claiming that we have more than one VEL gesture present in such segments. Browman & Goldstein (1986) suggest that oral and glottal gestures might also require a similar specification, a move which would at once double the number of gestures needed and which would result in an even more complex form of (12a). There is no or little evidence that such a move is necessary. Further, representations with two velic gestures are at odds with the conception that simple nasals and prenasals are differentiated by differences in the coordination of the velic and oral gestures. If on the other hand only prenasals need this type of representation, we are left with the problem of integrating a highly complex set of gestural relations which exist for one segment type only. What we need to capture is the fact that we have two separate but linked events - a period of nasality, which is accompanied by a period of overlapping (though not necessarily completely synchronous) supralaryngeal closure, which is followed by a period of non-nasal closure. There should be no need to differentiate between the types of nasal gestures found in simple nasals and prenasals, as Browman and Goldstein (1986) show for English. The fact that velic opening and closing may show differences in velocity and duration does not require that they should be expressed as separate gestures, only that a fuller specification in general is needed, as already noted. I will deal more fully with this issue in the following chapter.

Another problem arises with the representation of complex segments. (12b) clearly implies that both the LIPS and TB gestures are completely synchronous, overlapping each other completely. In AP, branching of association lines, unlike in FG as we have already seen, always implies temporal sequencing, thus avoiding the ambiguity allegedly present in the FG representations, and as we have no branching here, we have no temporal sequencing (assuming sequencing to refer to non-synchronicity in particular). As long as we have simple segments, i.e. segments with only one active oral gesture, the percolation principles of AP ensure that the correct value for CD is present in the Oral tube, and in general this can also be seen to hold true for complex segments.

Sagey (1986) showed that while a great many types of complex segments are found in the world's languages, they are all restricted to a single distinctive constriction degree, the non-distinctive CD being predictable.[13] A rather unusual form of representation is created by Sagey in order to match the manner degree with the appropriate articulator node, as can be seen in (15), representing Margi [pt], whose distinctive CD is **clo**, i.e. [-cont]. This is represented first of all at the top of the hierarchy, then a loop is drawn to connect this distinctive value with the appropriate articulator node, which in this instance is [coronal]. This is referred to as the major articulator. The manner degree of the minor articulator, being redundant, will be specified by the appropriate fill-in rules at a phonetic level - at the phonological level only the distinctive degree is specified.

(15)

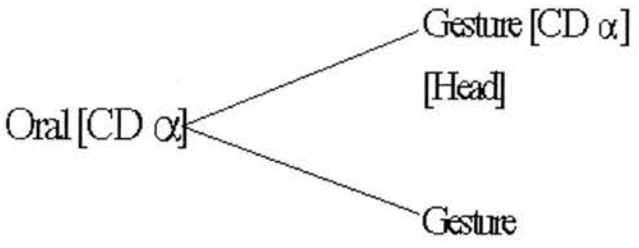


Basing themselves on Sagey, Browman & Goldstein suggest the representation in (16) as a possible AP equivalent. The major articulator is here marked as the head of a constellation of two oral gestures, and the head would by

[13]In fact, Padgett (1991) shows that at least for Kabardian we must specify more than one value underlying for [continuant], as for some complex segments neither of the values for [continuant] is predictable

convention automatically agree with the CD of the mother node (in this case the Oral node) and vice versa.[14]

(16)



As long as both gestures share the same CD, as in the example above from Margi, the interpretation of such a segment is straightforward. Problems arise, however, when the manner features for the major and minor articulators are not identical, as in [ps], or [kw]. Sagey claims that in Margi the head, or major articulator, for [ps] is the [+cont] portion, i.e. the [s]. This, if correct, poses severe problems for AP, as the percolation principles break down altogether here.[15] Given LIPS : **clo** and TT : **cri**, it is the tightest degree of constriction which would be percolated up to the Oral tube, i.e.. **clo**, and the coordination of the two gestures would result in [p] rather than [ps], as the lesser **cri** CD would be hidden by the tighter **clo** CD.

---

[14] There is some confusion here as to how precisely the CDs of the relevant gestures are to be specified. As we shall see, it is often unclear whether CD is an inherent property of the gesture, or is determined for it by some higher level. Browman & Goldstein claim that for complex segments, the normal percolation principles do not apply, except in a 'negative way'. The ramifications of abandoning the percolation principles for this small subset of segment types are not fully considered.

[15] In fact, Sagey's claim that the coronal segment, and thus [+cont], is the head, can be challenged to some degree. The identification of coronal as head is determined by principles of patterning and underspecification, but Sagey also uses the evidence of prenasalisation as evidence to support her claim. She cites Hoffman (1963: 29) to the effect that only stops and affricates can be prenasalised in Margi, i.e. segments whose distinctive CD is [-cont] - if [ps] is distinctively [+cont] then its non-participation in prenasalisation is predicted. She explains the apparent exception of [nl] by marking [l] as [-cont], despite its obvious continuant, fricative nature. However, forms such as 'msh', a prenasal labio-alveopalatal fricative, suggest that there are other exceptions showing that the choice of [+cont] rather than [-cont] as head has no phonological effect.

The only way in which the percolation principles could be preserved without making any further changes to the representation would be to assume that despite appearances to the contrary, the various articulators of complex segments were in fact ordered with respect to each other. Maddieson & Ladefoged (1989) indeed lend some support to this view. They show that for complex segments such as [gb] the labial and velar articulations are not in fact completely synchronous, a fact which the phonetic transcription itself seems to suggest. This they interpret as evidence that there is a clear difference between phonetic ordering and phonological unordering, but Browman & Goldstein claim instead that this lends support to their analysis. Given that there is in fact some ordering present, the choice between a theory which cannot show this and one which can seems fairly straightforward. However, a possible alternative would be to have a theory which could show ordering yet at the same time show that the gestures could act as a single unit for phonological rules.

At the simplest level, for a complex segment such as Margi /ps/, if the LIPS and TT gestures overlap completely then the percolation principles will allow only the tighter **clo** CD to filter up through the vocal tract hierarchy, and the allegedly distinctive frication would be completely hidden. If on the other hand the two gestures were offset with respect to each other to the extent that overlap were not complete, then both the period of silence/closure and the period of frication could be percolated correctly through the VTH (given appropriate reanalysis of the place of the VTH in the phonology). As the phonetic evidence of Maddieson & Ladefoged shows that the gestures are offset and do not overlap completely, we should reflect this fact in our representations. AP can percolate the correct values only if the two gestures are offset; the two gestures are indeed offset, therefore AP provides a correct formulation.

Presumably the necessity to include such timing relationships gives AP a representational advantage over FG, as AP not only remains consistent theory internally, but also more accurately reflects the phonetic facts. Unfortunately, there are a number of problems with this approach which make it less attractive. To begin with, if we look back at (12b) we see two gestures linked by non-branching association lines. By the conventions established by Browman & Goldstein, this diagram is to be read as referring to gestures which overlap completely, forming in the process a phonological unit. However, given the evidence from Maddieson & Ladefoged that the gestures are not completely synchronous, we would need some process to convert the synchronous sequencing of (12b) into a non-synchronous structure. Such a process would be completely ad hoc, and run counter to the general principles of AP in which diagrams are a direct reflection of the timing relationships present between gestures, with no need for an external mapping function.

If we abandon the Gestural Score of (12b) altogether in favour of the Vocal Tract Hierarchy of (16), we fare no better. We have no association lines telling us that the two gestures are phased with respect to each other in any way, so presumably no such relationship can be inferred, unless of course we subject it to the same principles as the gestural score in (12a). This, though, does not seem possible. The VTH for Browman & Goldstein does not contain any timing information, which is instead contained in the gestural score - the VTH is a snapshot in time and therefore all gestures should be interpreted as being simultaneous - and any number of gestures could be present with no implications for phonological ordering. The only way that we can read ordering from the diagram is to have access to information elsewhere, such that two active Oral gestures will always be slightly offset with respect to each other. This of course is exactly the condition necessary to interpret the FG representation in (12d), and thus would be subject to the same criticisms.

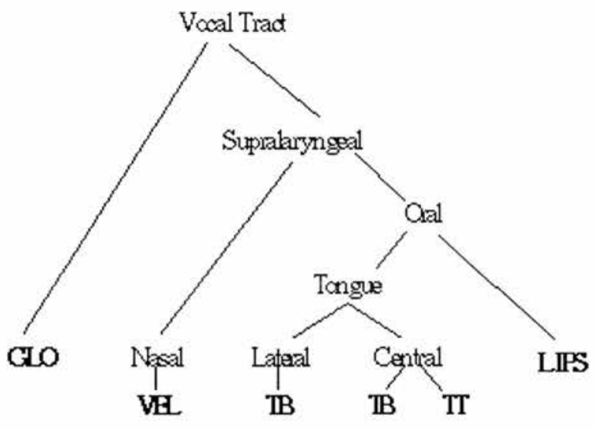
While it may be surprising to find AP and FG facing the same interpretative problems, the necessity to have access to information not contained in the diagrams is not necessarily a mark against a theory. Diagrams by definition require some knowledge of the world on the part of the reader, and are not in themselves phonological objects (Coleman & Local 1991). The problems only arise in FG when branching is used to describe very different events as in (12c,d). The tree structure of FG is not designed to include information on the relative timing of gestures, and Sagey's attempt to include this information without changing the basic structure results in the anomalies we have seen. Despite the criticisms which might fairly be made of the FG, it seems to me that the FG representations are in many ways superior to the AP representations as they stand. This is not meant to imply that the general direction of AP, and the criticisms which arise from it with regard to FG, are not valid, only that the particular conventions chosen by Browman & Goldstein themselves suffer from a number of problems which make them inadequate. Most pertinently, while claiming that ordering is essential both theory internally for AP and phonetically, AP is in fact unable to reflect this ordering in its representations. (16) above does not in any way imply ordering of the gestures, and while the principles of AP seem to be valid, the implementation is as yet inadequate. We need a richer but more constrained set of conventions.

Another set of problems arises when we try to discover precisely what is implied by the introduction of headed phonological structures into AP. Such structures are by no means uncommon in phonology, though they assume different levels of importance from theory to theory. On the one hand we find such as Dependency Phonology (Anderson and Ewen 1987) which, as the name suggests, is characterised by dependency, or government, at every level. On the other hand we have FG itself, where dependency plays only a minimal part. The question which needs to be asked here is whether Browman & Goldstein are justified in introducing such a concept into AP in order to handle one small class of segments, as there seems to be no evidence that it is to be extended to other or all segment types. If headedness accounts for only a single segment type, it functions less as a genuine relation between gestures and more as an arbitrary category marker. We must, then, consider what ramifications headed structures have for the theory as a whole.

To begin with, there now seem to be two different ways in which we can specify the CD of gestures. The basic principles of AP are based on the relative autonomy of gestures. As we have seen, apparently the only way in which the values of one gesture can affect the values of another is when the two gestures overlap. This can result in one of three things: hiding of one gesture by another so that only one is heard, as in 'perfec(t) memory'; blending of the values of the two gestures as in 'te<u>nth</u>'; or gestural undershoot, which is claimed to result in lenition. The introduction of headed structures is designed to avoid the first of these, hiding, segment internally. But what claims does this make for how gestures interact with each other and for the phonological status of the VTH?

If we look again at the VTH, given in (17), we see that the gestures themselves are all at the very bottom of the hierarchy and that the values present at the upper levels are entirely the product of the values of the gestures. At no point is this pattern reversed, i.e. the values of the gestures are never determined by higher levels of the hierarchy. Clearly, when we have blending or gestural diminution, the actual realised values of the gestures will be affected, but this should not be taken to automatically imply any change in the underlying values of any gesture - a gesture's underlying CD does not change simply because it is not achieved. At any event, this type of change is caused directly by the presence of another gesture, not by the VTH as such. Why, then, should we allow the Oral node to control the value of the TT gesture in /ps̱/?

(17)

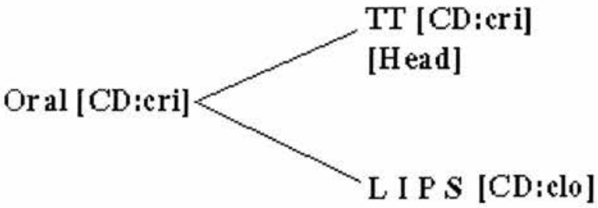


The TT gesture of /ps̱/ has its CD gesture constrained by the Oral node, when we would normally expect the Oral node to receive its value from the gestures below it, i.e. the TT gesture in tandem with the LIPS gesture. What then, if anything, governs the CD of the LIPS gesture? It would appear that while TT as head has some special constraining relationship with the Oral node, LIPS is governed by the normal conventions of AP, giving us two different ways of controlling gestures within a single segment.[16] However, despite the conflict created by allowing two seemingly incompatible methods of creating gestures to operate simultaneously within a single segment, this would not be enough to ensure the correct output, as is noted by Browman & Goldstein themselves. The main purpose then of the introduction of heads is to allow one and only one of the gestures in a complex segment to bear a *distinctive* CD, presumably leaving the redundant value of the other gesture to be specified by some other means. Again, we would need to ask where this specification would come from, would it be represented in the gestural score, and so on. The use of heads seems to be little more than an attempt to represent problems in FG that may not even be present in AP. Before introducing such ad hoc measures it is important to see if the same set of problems are actually present in AP.

There is another way of looking at the role of heads in AP. Whatever way we choose to represent distinctiveness and redundancy, whatever way we choose to generate the various values of gestures, for complex segments such as /p͡s/ we must ensure that the two gestures are offset in some way in order that both silence and frication are audible. This could allow us to view Browman & Goldstein's concept of headedness as implying ordering of some kind. Given two apparently simultaneous gestures, the presence of a head would mean that the gestures would need to be offset with respect to each other. Thus we could interpret (18) below as implying both ordering of the two gestures, and more specifically the fact that the head is ordered after the non-head. This allows us to avoid another problem inherent in (16), namely ensuring that we correctly percolate up the value of the non-distinctive CD. Without ordering, one and only one CD can be present at the Oral node.

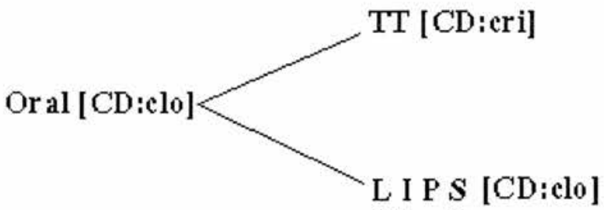
[16]Section 2.5 below discusses the suggestions of Bird (1991) where CD does not spread from any single point of the VTH, but is instead simultaneously specified for all points. However, as we shall see, this account of the VTH itself runs into problems.

(18)



The marking of TT as head for /ps/ as we have seen reflects Sagey's analysis, where the particular place of /ps/ in the pattern created by the Margi system suggested that the distinctive value for manner was [+cont] rather than the expected [-cont]. This is a reversal of the patterning usually claimed to hold in complex and contour segments, where it is typically the tightest constriction which has been claimed to be distinctive, for example in affricates or prenasal stops. In fact, as far as affricates are concerned it is doubtful whether either value for [cont] should be considered primary (Lombardi 1990). A further consequence of Sagey's claim is that we should find complex segments which are identical except for differences in which of the gestures is the head and which the non-head, so that we have for example /ps/ which has LIPS rather than TT as the head, as in (19). Sagey cites two contrasting labio-velar segments in Fula which become /g/ and /b/ respectively under a consonant gradation process, but the number of such segments might be reasonably expected to be higher across languages.

(19)



One difference we might expect as a result of changing heads is a change in the ordering of the two gestures. If the head were to be realised after the non-head, then (19) would represent the complex segment /s͜p/. This kind of argument is advocated by van de Weijer (1993), which I shall discuss in the following chapter. Following this line of argument, differences of ordering should be available for all complex segments. To a certain extent this is what we find. Recall the examples from Kru above, in which we find free variation between [gbe] ~ [bwe] ~ [gbwe], which would suggest that here we have the variation in ordering that Browman & Goldstein would predict, assuming either of the gestures can be a head. Unfortunately, it is clear that it is not simply phonetic variation that we would predict, but real phonological differences. Given labial and tongue body closure gestures, each of which can be a head as below, we would expect that a language could contrast (20a) and (20b), just as it would contrast the segments in (21), that is, we would expect contrasts to be maintained between /gb/ ~ /bg/ ~ /g/ ~ /b/. Of course, we do not find this situation in any language (Maddieson 1984). We could then say that given two active Oral gestures which act to form a single segment such as /gb/, if one of them can be a head the other can not, ruling out contrasts between e.g. /gb/ and /bg/, but again this would be a purely ad hoc and arbitrary measure.

(20)

(a)

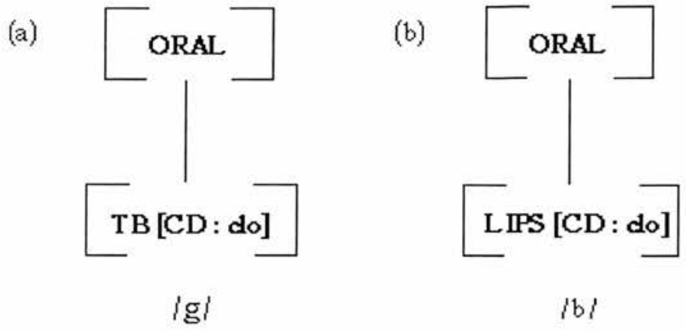
/gb/

Oral [CD:clo] — TB [CD:clo] [Head]
Oral [CD:clo] — LIPS [CD:clo]

(b)

Oral [CD:clo] — LIPS [CD:clo] [Head]
Oral [CD:clo] — TB [CD:clo]

(21)



Although these are relatively minor points, it seems as if the adoption of the FG representations by AP creates at least as many problems as it solves. What, then, are we left with? While there are clear difficulties with FG, there are equally large problems with AP. There is no coherent way to represent either contour or complex segments without creating confusion as to the presence or otherwise of ordering, and the introduction of heads seems to serve no purpose. While claiming that AP is 'inherently underspecified', Browman & Goldstein merely suggest that only active gestures are to be represented in the Gestural Score. The step from this to claiming that for complex segments such as /p͜s/ one of the gestures is not provided with a CD is an important one, and one which is not supported by any evidence within AP (as opposed to theory-internal evidence for FG). The introduction of redundancy and underspecification suggests that a greater number of contrasts, based on solely on differences in heads, should exist, yet there is little evidence to support this claim.

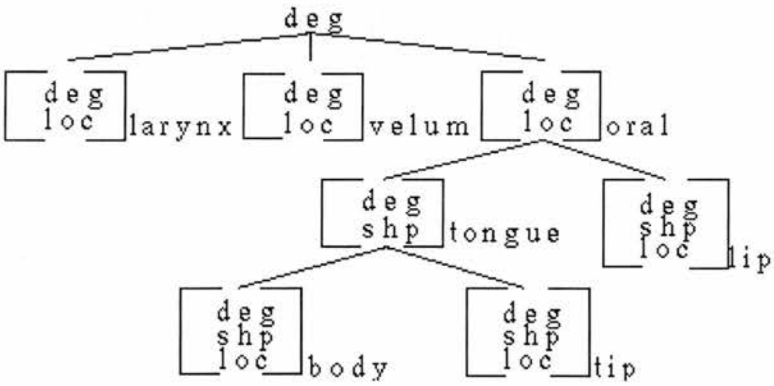
Many of the problems we have seen seem to stem from the reinterpretation of the VTH as somehow a licenser of gestures and gestural values rather than a product of those same gestures. While gestures remain independent of the VTH these interpretative problems do not exist, yet without some form of hierarchical structure it seems impossible to capture the fact that most complex segments, if not all, do seem to show predictability in their values for [cont]. In the following section I discuss the proposals of Bird (1990) who attempts to show that the problems raised by Sagey are amenable to analysis in AP terms.

2.5

Bird's (1990) Approach to AP

Bird (1990) shows that it is possible to solve some of the problems of AP by introducing the conventions of Sagey (1986) in a more coherent fashion. The basis for the synthesis is a time and event based logic, but what is important for us here is the manner in which Bird chooses to represent headedness and percolation. Bird's geometry, shown in (22) below, closely follows AP's in most details, though there are a number of minor differences. For example, the larynx (GLO in AP) is provided with a feature for location, in order to represent glottal raising and lowering, which should also be incorporated into AP, as Browman & Goldstein (1992a) themselves comment. More importantly, however, Bird follows AP in placing degree of closure on each and every node, so that it is present at all levels of the tree.

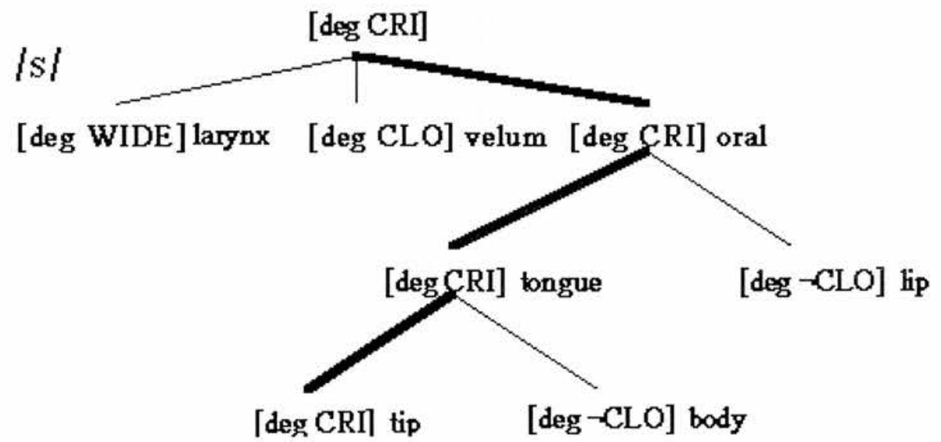
(22)



All of the nodes, except the Oral node, correspond directly to articulators, but this should not be taken to imply that Bird is advocating a gestural approach directly akin to AP. Instead, both Bird and AP share the view that the simplest way to model a feature hierarchy is to assume that it arises from the fact that all speakers of all languages share what is essentially the same vocal apparatus. While this is compatible with a gestural approach, Bird is more concerned with the theoretical aspect of the tree, rather than with directly modelling physical processes. Having said this, the principles of interpretation used by Bird are closely modelled on Tube Geometry, and can only be understood with reference to it.

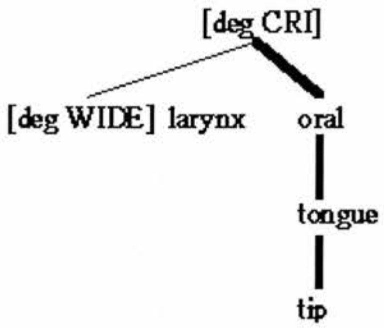
As in AP and FG, it is the nodes below the topmost [degree] node (which Bird refers to as the root node), which control the value of the root node itself. Thus, a critical degree of constriction at the tongue body should percolate upwards through the tree, all things being equal, to give a critical value at the root node. The only constraints placed upon this percolation are that the value of CD at any node be constrained by the values of the daughter and mother nodes. In (23) we see a representation of /s/. Here Bird introduces a simple system of marking the major articulator, and at the same time obviates the need for directionality of percolation. The [**cri**] CD is marked at each level of the tree, from the root to the tongue tip, yet is not confined to any of these levels, and can not be said to begin at either the root or the tongue tip. Rather, the solid black lines are to be seen as part of the structure, not simply a handy, if ad hoc, notation like the pointer system of Sagey. They actually specify the path that percolation must take, thereby linking the top and bottom of the tree. The interpretation of this is non-directional, so the value of each node is dependent on the values of all the other nodes.

(23)



In (23) the values for the lip and tongue body nodes are not fully specified. Instead, they are both specified as [¬**clo**], a negative specification which allows any value to be chosen except full closure. This information can be omitted with no loss of clarity, as in (24), which contains the same information as in the Gestural Score seen in (7) above, that is it contains only unpredictable information (Bird assumes the same default settings for gestures as Browman & Goldstein). What, though, do the emboldened lines mean?

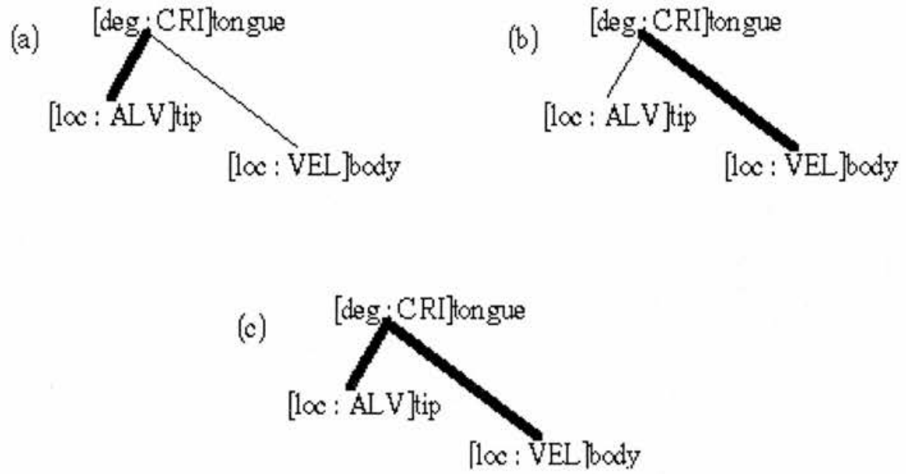
(24)



Just as each node is part of the overall tree, so it is also part of a smaller, local tree, and as such it is either the head of that tree or a non-head. There are no other relationships. A node which is joined to a higher or parent node by an emboldened line is the head constituent of its particular local tree, and nodes which are not connected in this way (i.e. those joined to the parent node by a non-emboldened line) are non-head constituents. For example, the tongue tip node in (23) is the head of its local tree, with the tongue body node as a non-head constituent and the tongue node as the parent node. The head node is constrained to have the same CD as its parent node, and vice versa, and this must be compatible with the percolation principles of the theory, which are those of Tube Geometry. This is a crucial point, to which I shall return in a moment.

One of the major advantages of this system is that it we need only refer to the degree of constriction at one point in the tree, but at the same time we are aware that this same CD is repeated at a number of nodes, without having to stipulate this. In (25), taken from Bird (1990), we see representations for three different fricatives. (25a-b) could give /s/ and /x/ respectively, while (25c) is a possible representation for the complex Swedish segment 'kj', a doubly articulated voiceless fricative. Note how the head of each tree automatically agrees with the mother node in CD, while in (25c), where there are two head nodes, both agree with the mother node. How then does this compare with the AP forms, and what advantages, if any, does it give us?

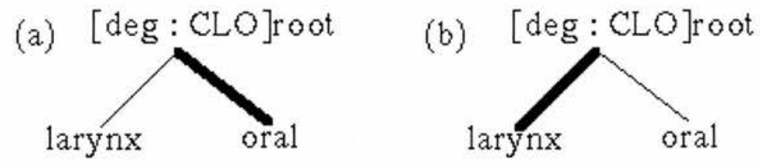
(25)



At first sight it appears that we do indeed only need to stipulate degree of closure once, but is this really the case? In AP, it is the gestures at the base of the tree which control the CD at higher levels, and so the only time at which CD need be specified is at the gestural level - the value of CD at higher levels of the tree is a function of the percolation principles of Tube Geometry, just as for Bird. This holds for all levels, including the Vocal Tract level. FG works along similar lines, with the root node being merely a focal point to which the features below it can attach. For Bird, however, the root node appears to be much more, acting much like a licenser (in the sense of Goldsmith 1991). As it is at the root level that the major CD is located it seems reasonable to assume that it is this node which licenses the CD at all other nodes which it dominates, even though percolation is assumed to be equational and thus non-directional. Locating the CD at a lower level of the tree could lead to incorrect percolation (note that there are no segments in which an emboldened line is not connected to the Root node).

A further problem also affects the way we must view the root node's function. (26) (from Bird 1990) shows another way in which percolation can work. (26a) can be viewed as the default mechanism for percolation, down from the root node to the bottom of the tree. The default direction would be through the oral node, leaving the larynx to receive its value in some unspecified way. In appropriate circumstances, however, the default direction might be overruled in favour of some other system, in this case the CD of the root node being percolated to the laryngeal node, presumably leaving the oral node to receive some other, default value, or no value at all. On this interpretation, root is more than a convenient docking site for the degree of constriction, it actually licenses the occurrence of a specific value and controls the direction of percolation. This would provide an appropriate description for such processes as the interchange between /t/ and /ʔ/. This, however, will have a substantial effect on our view of the role of the root node, for not only must we provide a CD for the major articulator (tongue tip in (24)) but we must also provide an overall CD for the whole tree, which will be specified at the root node. For example, the alveolar nasal /n/ will have a tongue tip CD of **clo**, which will be specified at the root node, but which will be simultaneously present at the root, oral, tongue and tongue tip nodes. In addition, the velum node has a CD of **open**.[17]

(26)



[17] Quite where and how this is specified, and how it differs from the way in which the tongue tip node receives its CD, is not relevant here, though it is an interesting point.

As said earlier, Bird employs the percolation principles of Tube Geometry, but the different arrangement of his tree results in percolation working in a slightly different manner. Bird does not provide a lateral tube, preferring instead to specify laterals by using the feature [shape], and as a result the tongue tip and tongue body tubes are linked in series, as well as the tongue and lip tubes. The oral and nasal tubes are then combined in parallel (though this is a point to which I shall return), the combined oral-nasal tube then being joined in series with the larynx to give the overall value of the tree. As in AP, the overall value of a nasal is **open**, but this gives rise to ambiguity as to the value of the root node. While it is **clo** in accordance with the emboldened line notation and the percolation principles attached to it, the overall value of the tree, expressed in the CD of the root node, is **open**. Thus root is simultaneously **open** and **clo**, a paradox which can not be solved. Further, this can not be resolved no matter which view of percolation we take, the downward view shown in (26), or the non-directional one of (23).

The root node is asked to play two incompatible roles by Bird. In a sense it must act as both licenser and licensee, and as long as there is no conflict in values there is no problem. Problems only arise with a small number of segment types, but the presence of these suggests that we would be wise to restrict the role of the root node in some way, so that it fulfils one or the other of these roles, but not both. In addition, it seems unsatisfactory to say the least to allow the root node to supply oral nodes with their CD, but to have no similar mechanism for specifying the CD of the velum and larynx nodes. In /n/ the CD of the velum node is apparently specified by quite different means from the CD of the tongue tip node, with no explanation as to what this difference is or why it exists.

There is a further problem with Bird's theory, again involving the oral and nasal nodes. Bird invokes the arguments of McCarthy (1988) and Iverson (1989) to remove the need for a Supralaryngeal node, his tree resembling the input tree of AP rather than the output. Tube Geometry, as described in Chapter 1, is modelled on the passage of air through complex tubes, and for this purpose the oral-nasal tube, or Supralaryngeal tube, is vital. The use of Tube Geometry in AP demands that the

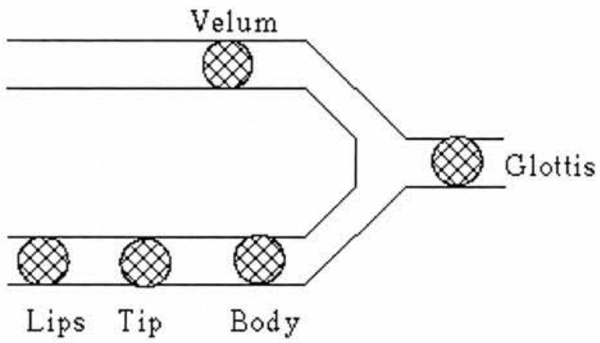
Supralaryngeal node (or complex tube) is as real a node as the Oral or Tongue Body nodes. The same must therefore be true of the use of Tube Geometry in Bird's theory, at least in terms of calculating the overall value of the VTH - if the oral-nasal configuration were not a tube then how could we apply the principles of Tube Geometry to it? If we use Tube Geometry, then the Supralaryngeal tube must be present, if it is not then we cannot use Tube Geometry.

Let us assume then that we do have a Supralaryngeal tube. Given that we know its value is the minimum[18] of the combined value of the gestures, we know that the oral and nasal tubes are joined in parallel, as should be clear from (27), adapted from Bird (1990), which is a simplified tube model of the vocal tract. While the head nodes must by definition agree in CD with their mother nodes, this must be compatible with the percolation principles, as we have seen. As an example of this, Bird states that if we have a complex node, the constituents of which are joined in parallel, such as the Supralaryngeal node, then the head constituent must be *at least as constricted* as all the other constituents. How does this apply here? For the Supralaryngeal node it is the Oral node which will be the Head. If we have for example a stop such as /t/, with Oral : **clo** and Velum : **clo** we have no problems. For a nasal fricative, e.g. /ṽ/, we have Oral : **cri** and Velum : **open**, which again presents no problems. For a non-nasal fricative such as [s], however, we have Oral : **cri** and Velum : **clo**. This is problematical on two fronts. First of all, the non-nasal fricatives break Bird's restrictions on possible values, since the non-head has a more constricted CD than the head, thus breaking the percolation principles. Secondly, we have an easy formulation for nasal fricatives and no such formulation for oral fricatives, the reverse of the situation we would wish. By moving away from a model in which constriction degree of the overall vocal tract is strictly a function of the interaction of the constriction degree of the active gestures, towards one in which the CD is predicated of the vocal tract as a whole, we create this serious anomaly in

[18] Recalling that we are using Bird's definition of minimum and maximum, as discussed in Chapter 1, in which minimum refers to the least constriction degree, and maximum to the maximum, as opposed to the usage of Browman & Goldstein (1989).

which it is easy to describe a very rare segment type, and impossible to describe a very common one.

(27)



2.6
Conclusions

Phonology is distinct from phonetics, it is argued, by virtue of its containing categorical structures and processes in distinction to gradient structures and processes. Many of the problems facing AP stem directly from its attempts to bridge the gaps between these two domains. AP has successfully challenged many of the assumptions of phonology as to what is and is not a categorical process, reducing many phenomena previously thought to be discrete to the automatic results of gestural overlap and deletion. AP's position is greatly weakened, however, in that for all that it gives the appearance of being a phonology, once we scratch the surface there appears to be little in the way of categorical structures, and worse still is the fact that what categorical structures there are appear to make incorrect predictions about a wide range of phenomena.

The problems of the categorical structures so far proposed for AP stem from two main areas. Firstly, the respective roles of the various components of AP - the Vocal Tract Hierarchy, Gestural Score, gestural overlap etc. - are at best unclear and at worst confusing and contradictory. For aspirated stops, for example, if we base our observations on the gestures within the Gestural Score we gain a clear picture of the timing relationships between the two gestures, but nowhere in the Vocal Tract Hierarchy are the relationships between the component gestures made clear. In addition, the change from complete overlap for /p/ to partial overlap for /p^h/ and minimal overlap for /p + h/, while perhaps valid from the perspective of quantal theory, is nevertheless an arbitrary one given both the many additional types of overlap possible between even two gestures as noted by Clements (1992) and the failure to group /p/ and /p^h/ together as single segments as distinct from the bisegmental /p + h/.

Secondly, the root cause of many of the problems in the structures we have seen in this chapter can be found in the use of the structures of FG as the departure point. As we have seen, FG has been very successful in describing many areas of phonology, though of course it has been more successful in some areas than in others, but the primitives of the theory - that is, features - are very different in nature to those of AP. One area in which the two theories do appear to show some convergence, however, is in the Vocal Tract Hierarchy of AP and the various trees of FG, and it is this similarity which has perhaps led to attempts to adopt other aspects of the structures of FG directly into AP. The result of this has been both positive and negative.

For example, Browman & Goldstein's headed structures, mimicking those of FG, fail in many aspects to adequately describe the range of segment types found in natural languages. On the other hand, Bird's synthesis of Sagey's pointer notation and AP shows that there are alternative ways of incorporating both the timing dimension of AP, and its physically based feature hierarchy. However, while it manages to avoid many of the defects of Browman & Goldstein's approach, it does have defects of its own while at the same time providing some simple and elegant solutions to a number of problems.

Rather than attempting to incorporate the structures of FG into a gestural-based model, we may find more success in deriving a model independently of FG. In the following chapter, therefore, I will suggest other directions which might be profitably explored in order to answer some of the questions raised here.

CHAPTER 3
SEGMENTS IN ARTICULATORY PHONOLOGY

3
Introduction

AP suffers, if that is the correct word, from its lack of restrictions on both the number and type of relationships it allows to exist between gestures. The criticisms of Clements (1992) and others still stand, that AP appears to allow a far higher number of phasing relationships to hold between gestures than can be shown to be phonologically useful or necessary. Although it may be true that a larger number of relationships might be useful than these criticisms might suggest (cf. Steriade 1991), no distinctions are drawn between the functions of these relationships, leading to the creation of false natural classes.

In this chapter, I present a theory of headed structures which will provide AP with the necessary categorical representations to distinguish between such as /nd/ and /n + d/. While maintaining the use of heads as introduced by Browman & Goldstein (1989) I interpret them in a very different way. The chapter is laid out as follows: In **3.1**, I discuss the general principles of gestural coordination in terms of segmental structure; **3.2** discusses the view of segments as containing headed articulatory structures with audio-acoustic goals, such that all segment types are distinguished only in terms of whether or not their component gestures are heads; in **3.3** I discuss the respective roles of categorical and gradient information in the realisation of these headed structures, in particular the role of the glottis in nasal segments; in **3.4** I compare the respective predictions of single and two Root theories of Feature Geometry with those of AP regarding prenasalised and complex segments; finally, in **3.5** I discuss the representation of affricates.

3.1
The Goals of Gestural Coordination

The question still remains as to how precisely we are to make the different patterns of gestural overlap required by AP distinctive so that we may move towards categorical structures, though a question of this type is not restricted to this theory alone. The recognition that not all the component parts of a segment are (phonetically) simultaneous leads us to the search for a mechanism to coordinate these various parts, a task which confronts all phonological theories, as although absolute simultaneity is very much the exception rather than the rule, two or more gestures can behave as if they were simultaneous to form a single unit; in other words, we have a distinction between phonetic ordering and phonological unordering. AP and Feature Geometry both agree in breaking the vocal tract down into a number of discrete units, and while they might not agree on the identity or interrelationships of these units, both recognise that they combine to form a hierarchy of some kind. In other words, although we might have separate gestures or features, we view the vocal tract as being an integrated whole (Ohala 1990). This is especially important for AP because all such ordering and coordination of gestures must be directly incorporated into the phonology, as we saw in chapter 2. While each gesture or feature, therefore, has a degree of independence, if we wish to maintain the idea of a segment we must provide some non-arbitrary means of coordinating these independent units to produce exactly the right combination. If we consider again the case of voiceless aspirated and unaspirated stops, we know that the gestures for both segments are identical, and that the two sets of stops differ only in the relationships that hold between the gestures. We know, or think we know, what these relationships are, but what we do not know is how or why these relationships hold. This is the question to be answered here.

Despite AP's basis in articulation, it would be a mistake to assume that it is solely articulatory in nature. As both Clements (1992) and Kohler (1992) have noted, by basing the various constriction degrees on quantal theory AP implicitly accepts that while speech production might very well be capable of being expressed largely in articulatory terms, we speak in order to be heard and understood (Jakobson et al. 1951). In other words, the articulatory gestures of AP would be of no use as a communication tool unless they also produced specific acoustic and auditory events. It is on this same basis that Kingston (1990) attempts to provide an answer to the problem of the internal coordination of the component gestures of segments.

Kingston takes as his starting point the nature of the relationship between glottal and oral gestures. The most salient point of a stop, he argues, is its release burst, containing most of the cues as to the nature of the stop. The role of the glottis in this burst is in many ways as important as the place of constriction, exemplified by the fact that many languages contrast stops not only by place but also by the state of the glottis. If we take a simple opposition, such as that between /p/ and /b/, we can see what Kingston refers to as the 'proximate' role of the glottis, that is its effect on source characteristics, in this case the creation of a contrast of voicelessness versus voicing. Kingston claims that in addition to this proximate role the glottis also has a secondary or 'distal' role in its effect on the nature of the release burst. In a contrast between voiceless aspirated and unaspirated stops, it becomes crucially important just when the glottis achieves its maximum degree of opening, for while it is the closure of the stop that initially creates the pressure build up which leads to the burst, it is the glottis which controls the rate and manner of airflow, thus controlling the nature of the pressure build up and by extension the character of the release burst.

This role of the glottis is confined to stops. Approximants show no such behaviour, it making little difference quite when or how accompanying glottal gestures are coordinated,[1] and while for fricatives the glottis is typically wider than for other segments, it has no distal function as such and any glottal gesture is generally evenly spread throughout the period of frication. If, then, the glottis has such a function when allied to the oral closure gesture of a stop, Kingston argues that it is reasonable to assume that as it is the release of the closure which is affected,

[1]As approximants are mostly spontaneously voiced, it is debatable whether e.g. a segments such as /l/ should have an active glottal gesture at all.

glottal gestures should be constrained to phase with this release portion.[2] The narrower the oral constriction, the tighter the timing of the glottal gesture is 'bound', so that for stops there will be a greater positive correlation between the glottal gesture and the release of the closure than between the glottal gesture and the onset of closure. For other types of constriction the glottis has relatively greater freedom in its precise timing, so that for example a glottalised lateral may be realised with glottal closure before, during or after its accompanying oral gestures, or a combination of these, as the overall acoustic effects of each realisation will differ very little (Goldstein 1990).

Unfortunately, Kingston found that the binding hypothesis fails in a number of ways. Given a fricative-stop cluster such as /st/ in English 'sting', where a single glottal opening gesture is shared between the two consonants, we would expect the glottal gesture to be dominated by the release of closure for the stop and to continue to positively correlate with it, but in fact it is the fricative which seems to dominate. Alone, peak glottal opening would occur around the midpoint of /s/, and while there is some compromise (peak glottal opening for /st/ occurs at the 'border' of the two consonants) it is clear that the glottal gesture is not phased with the oral release of the stop. Again, for Icelandic the glottal gesture of postaspirated stops in stop-vowel sequences appears to correlate positively with the following vowel rather than with the accompanying oral closure, casting further doubt on the generality of the binding theory.

However, while the stricter interpretation of Kingston's theory does not appear to hold, the general principle is undoubtedly correct, and Ohala (1990) proposes a possible compromise solution based upon Kingston's earlier study on the differing behaviour of glottalised sonorants and stops (Kingston 1985). The glottal closing gesture for a segment such as /m'/ is relatively free as to its coordination with the supralaryngeal gestures and may spread to adjacent segments, so that it is

[2]Steriade (1992) makes similar claims. She suggests that stops are distinguished from other segments in that they consist of two ordered parts, namely closure and release, to which other features such as voicing can link.

difficult to know whether we should transcribe such a sound phonetically as [ʔm] or [mʔ]. The same gesture for glottalised stops such as ejectives is, in contrast, much more tightly constrained and considerably less free to spread to adjacent segments. Ohala suggests that rather than employing strict coordination between glottal and oral gestures, a requirement of simple overlap might be sufficient. This does not imply that we can simply avoid specifying the precise coordination holding between gestures, as languages do differ as to their coordinative patterns, but what it does mean is that the goals of these coordinations would be very much simpler than those suggested by Kingston. What is important is, in Ohala's terms, that there be the correct 'cooccurrence of states' (p. 436).

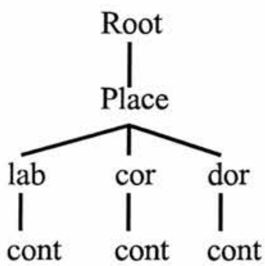
To return to the case of aspirated and unaspirated stops, we recall that in AP they are identical in terms of their component gestures, differing only in the way in which these gestures are coordinated. There are two assumptions that have generally been made regarding these different phasing relationships. First of all, we assume that each gesture consists of three different points which are available for coordination - onset, target and offset. When we phase two gestures with each other, what we actually phase are one of the three points of one gesture with one of the corresponding points of another. Secondly, we recognise three different types of overlap which form a semi-abstract series - complete, partial and minimal. The basis for this three way distinction is clearly quantal, assuming that there are only three distinct ways in which two gestures can coordinate, but as Clements (1992) points out, as soon as we introduce a third, or even a fourth gesture the number of possibilities increases exponentially, creating a far greater number of contrasts than is necessary for any phonology. Browman & Goldstein (1992b) suggest that at some level this larger number of contrasts might be needed, though the extent to which such contrasts could be phonologically useful is unclear; but what does seem clear is that we need some way to constrain both the number and the type of gestural relationships to produce a much smaller set of possible outcomes.

The incorporation of quantal theory into AP produces a blend of articulatory and audio-acoustic data, though the primary concern of AP when coordinating gestures remains articulation. However, what Ohala makes clear is that it is not simply or even primarily the actual physical coordination of gestures which is important, but rather what the coordination produces. If we assume that what is important about the distinction between complete, partial and minimal overlap is not the overlap itself but what it results in, then we can see the distinction between e.g. /p/ and /p^h/ in a new light. What is important for /p/ is that there be a period of voiceless closure, but what is important for /p^h/ is not only that the period of closure be voiceless, but that upon release of the closure this voicelessness continue as postaspiration. In other words, whereas /p/ consists of two gestures producing a single event of voiceless closure, /p^h/ consists of the same two gestures producing two events, voiceless closure followed by voiceless noise, and I suggest that it is the creation of these ordered events which is the sole goal of the phasing relationships of AP. What we must now determine is what these events are and how we can construct a phonology which refers to them. It is to these questions which I now turn.

3.2
The Internal Structure of Segments
3.2.1
Dependency Structures in AP

Standard feature geometry gives equal weight to all features of the tree, whatever their position. However, not all relationships between features are the same, as discussed in chapter 2. Features such as [place] and [laryngeal] feed directly into the top of the tree, the root node, but each of these features can dominate individual features and even trees themselves. A feature such as [coronal], while in many ways equal to features such as [cont] or [laryngeal], in that for example all can spread, is nevertheless a stage further removed from the top of the tree by virtue of its being dominated by an intermediate node, [place]. In Padgett (1991) [cont] is still further removed from the root node, being a dependent of all place nodes, so that for e.g. /s/ we have a relationship [place] > [coronal] > [+ cont] as we can see in (1) below. In terms of the structures of Sagey (1986) or Bird (1990), discussed in chapter 2, [place] functions as a head, dominating a dependent [coronal] which in turn is also a head, governing a dependent [+cont].[3] The chief consequence of such a structure is that while the dependent can spread independently of the head, spreading of the head automatically entails spreading of the dependent. A [cont] node dependent on a [place] node is thus predicted to behave rather differently from one which is directly dependent on the Root node, but outside of spreading and deletion operations based on the hierarchical implications, the relationship of dependency plays little part in the theory (see McCarthy 1988).

(1)



This type of headedness is of course also present in the tree structure of AP, though since feature geometry is not bound by the anatomical structure of AP it has much more freedom in its choice of heads and dependants (cf. Piggot 1992). There are, however, different interpretations possible of the role and function of heads in phonology, the chief of these being to give heads greater 'prominence' than non-heads. This is characteristic of both Dependency Phonology (DP) (Anderson & Ewen 1987) and Government Phonology (e.g. Harris 1990; Kaye et al 1990), though the precise realisation of prominence differs in the two. The description given in DP of aspirated and unaspirated stops is particularly relevant here. I base the following discussion primarily on Anderson & Ewen (1987).

DP is made up of a number of gestures - though of a very different type to those of AP - and these gestures in turn are composed of a number of components.

---

[3] By referring to [Place] as a head, I imply nothing more than that it may have another node subordinate to it (i.e. dependent on it).

These components can form phonological relationships with other components within the same gesture. In addition, gestures themselves can combine to form phonological relationships with other gestures. The basic gesture used to distinguish manner is the Categorical gesture, which is composed of two different components, |C| and |V|. |C| is defined as a component of 'periodic energy reduction' and |V| as 'relatively 'periodic', and when standing alone i.e. when not combined with any other gesture or component, represent a voiceless stop[4] and a vowel respectively. |C| and |V| can also combine with each other in a number of different ways.

Alone |C| and |V| stand at opposite ends of a hierarchy, but in combination they can form a continuous chain along that hierarchy, moving from stop to fricative to approximant and on to true vowel. For example, if we were to combine a single |C| with a single |V|, there are three possibilities, given in tree form in (2) along with non-combined |C| and |V| (where |u| is a component from the Articulatory gesture specifying labiality). The component at the top of the tree, the head, is more prominent than the non-head or dependent which it governs, and as such the property characteristic of the head is that much greater, and the property characteristic of the dependent is that much less. In (2a) the simple |C| is equivalent to a voiceless stop /p/, but the addition of a dependent |V| in (2b) forms a segment which is one degree more periodic, i.e. /b/. By reversing the dependency relationship in (2d) the prominence of the periodicity is that much greater, and the result is the nasal /m/.[5] Differences in the identity of the head thus translate directly into differences in segment type. However, it is constructions such as (2c) which are the most interesting from our viewpoint.

---

[4]|C| represents an oral stop only when combined with a specification for place, otherwise it represents a glottal stop.

[5]In fact, we would require in addition a component of nasality for /m/, not shown here.

(2)

a.	b.	c.	d.	e.
C	C	V:C	V	V
	|		|	
	V		C	
{u}	{u}	{u}	{u}	{u}
/p/	/b/	/f/	/m/	/u/

(2c) introduces the notion of mutual dependency, where both components are simultaneously head and dependent of each other, resulting in a voiceless fricative /f/. Although the possibility that phonological structures can consist of more than one head is not unique to DP, it is characteristic of it, and as we shall see below it is directly transferable to AP, though with a different interpretation. This three way dependency - head, mutual and dependent - is crucial to the DP analysis of aspirates.

Korean (Kim 1970) has a three way contrast in its stop series, namely aspirated, unaspirated, and tense or glottalised, illustrated in (3), where series I is tense i.e. glottalised, series II unaspirated and series III aspirated. These are shown in DP notation in (4), where a component of 'glottal opening', |O|, is introduced, based on Kim's correlation between aspiration and glottal opening. What is important here is that DP is able to describe the stop series in terms of precisely the same primitives simply by specifying different relationships between the various components, without recourse to such additional features as [spread glottis].[6] As we have already seen, voiceless aspirated and unaspirated stops in AP likewise consist of the same gestures, so the question arises as to whether it is possible to describe the different gestural relationships of AP in a similar way to that of Dependency Phonology.

[6]There is some argument in the literature over whether the glottal opening gesture should be in the Categorial gesture, or in a separate gesture of its own (Ewen 1982), but this does not affect the discussion here.

(3)

a.	/p'ali/ 'washer'	/t'al/ 'daughter'	/k'ali/ 'villain'
b.	/pal/ 'leg'	/tal/ 'moon'	/kali/ 'stack'
c.	/p^hal/ 'arm'	/t^hal/ 'mistake'	/k^hal/ 'knife'

(4)

C	O:C	O
|		|
O		C
/p'/	/p/	/p^h/

Let us again take as our examples the voiceless unaspirated stop /p/ and the voiceless aspirated stop /p^h/. In gestural terms both of these segments contain the same two gestures, namely LIPS : **clo**, and GLO : **open**; the fact that in each case the gestures are to all intents and purposes identical means that the differences in what they produce - i.e. /p/ in one case and /p^h/ in the other - must lie not in the gestures themselves but in the relationships between them. It is these relationships which we must describe. However, in order to fully understand the behaviour of the differences in these gestures when in combination with each other or with other gestures, we should first examine their behaviour when they are not so combined; that is, when they stand alone. This is particularly relevant given the fact that both of the constituent gestures of /p/ and /p^h/ are themselves able to form segments without needing to be combined with any other gesture, so that a LIPS : **clo** alone is identified as the segment /b/, while an unaccompanied GLO : **open** gesture forms /h/.

Clearly, a single gesture such as LIPS : **clo** does not by itself provide all of the components necessary for a full description of /b/. Along with the single gesture for /b/ there are, as we saw in chapter 1, a set of accompanying neutral settings for each articulator so that while for /b/ it may be true to claim that LIPS : **clo** is the sole gesture, it would be inaccurate to say that it is the sole component. For example, the neutral setting of the glottis as critical ensures that /b/ is characterised by its being voiced, while the setting of the velum as raised prevents any venting of the airflow through it. All of these factors combine in the formation of /b/, and a change in any of the parameters would result in the creation of something other than /b/. What, then, is the role of the LIPS gesture in /b/?

In the production of speech, three separate components are generally identified as being of particular importance: initiation, articulation, and phonation (Catford 1977). The first of these, initiation, provides the necessary flow of air which is then modified to produce the various sounds of speech; no speech is possible without it. However, while initiation is clearly of fundamental importance, it will not be the focus of discussion in this work, other than to note that all of the segment types to be discussed are characterised by pulmonic egressive airflow. We shall instead concentrate on the ways in which this airflow is affected by the two remaining components, articulation and phonation. Articulation can be defined as a modification or shaping of the initiatory airflow by the various articulators in the vocal tract which generates sounds of specific types, and like initiation it is an obligatory component i.e. without it no speech is possible. The third component, phonation, can be defined as the creation of audible acoustic energy at the larynx, but in order for this energy to be created there must be an appropriate airstream through the larynx.[7] However, the existence of segments such as glottal stop [ʔ], where there is complete closure of the glottis and thus no airflow and hence no possibility of phonation occurring, shows that while all languages employ phonation in some way, unlike initiation and articulation it is not an obligatory component for all segment types.

[7] In fact, there is some conflict between acoustic and articulatory definitions of phonation, but these do not affect the argument (see Ladefoged & Maddieson (1996) for discussion of this point).

The same absence of phonation is found in /h/, which in gestural terms can be straightforwardly described as a GLO : **open** gesture, the glottal gesture acting as an articulator upon the airflow, not as a phonator (Catford 1977). In contrast to /h/, the voiced stop /b/, as its description suggests, does involve a phonatory component, namely modal voicing, but this results not from any gestural activity of the glottis but directly from the glottis's neutral setting. Both /b/ and /h/, then, are described in AP as containing single gestures which in each case correspond to a single segment as defined in chapter 2, and in both cases the gesture also describes solely articulatory information, any phonatory information being accounted for without reference to gestures. In fact, in all instances when a segment contains a single gesture, that gesture must always be considered as articulation; the role of the LIPS and GLO gestures for both /b/ and /h/ thus suggests that amongst the various components of speech, articulation has particular prominence.

This articulatory role of the gestures in /b/ and /h/ must derive from their role in the phonology, and should therefore be included in any description of the internal structure of the segments. In addition, and in contrast to the structures examined in chapter 2, it should be assumed in attempting to provide such a description that it is not only complex segments such as /k͜p/ or /p͜s/ which require a more elaborate internal structure involving heads and non-heads, but that all segments must fulfil the same basic structural requirements. By making this assumption we at the very least avoid some of the interpretative problems of the forms discussed in the previous chapter where heads were only required for a subset of segment types. So, rather than assuming that only a very restricted number and type of segments are best described in terms of headed structures, I instead assume that all segments should be described in these terms, i.e. each segment must contain a head.

Given that both /h/ and /b/ each contain only single gestures, together with the straightforward assumption that only gestures can be heads, it naturally follows that in each case it is these gestures which must be the heads, GLO : **open** in the case of /h/ and LIPS : **clo** in the case of /b/; there are no other gestures available. The resulting structures are shown in (5a,b), where the superscript $\{^{H}\}$ indicates headedness. There is no need for any equivalent of the root node of FG in order to mark the structures in (5) as segments. Instead, as in the DP representations in (2) above, the marking of e.g. the LIPS : **clo** gesture as a head is sufficient to encode its segmental status without recourse to any higher level nodes. In other words, /b/ is a unit which contains a LIPS : **clo** gesture and no other gestures, where the LIPS gesture is a head. Similarly, /h/ is a unit which contains only a headed GLO : **open** gesture.

(5)

a) GLO^{H} : **open**

b) LIPS^{H} : **clo**

In terms of the phonological structure of /b/ and /h/, (5a,b) are complete descriptions; there is no need for any further hierarchical structure and no need to refer to the vocal tract hierarchy (VTH) or the gestural score. However, while the reasoning above is perhaps adequate to identify the gestures in (5) as heads, matters are not so straightforward for segments containing two or more gestures where we require some other means of determining what does and does not constitute a head. Characteristic of heads, as we saw above in the discussion of DP, is their greater prominence relative to non-heads. If the heads of AP were similarly to exhibit greater prominence, the physical nature of gestures would suggest that this prominence be realised in terms of the physical relationships between heads and the rest of the vocal system. Following this reasoning, I propose that heads in AP are subject to the principles in (6), where prominence is interpreted in physical terms.

(6) Head Identification in AP

a) A gesture is a head if it is true that there is a period during its target portion when it alone dominates the vocal tract.
b) A gesture dominates the vocal tract if it: (i) controls the neutral settings of the various articulators; (ii) controls all other active gestures within the segment.

What, though, are the implications of these principles for gestures? As we saw in chapter 1, when an articulator is not involved in gestural production it adopts a neutral position, so that e.g. the lips take up a neutral spread open position for /d/. A similar neutral position is given for each articulator (see diagram (11) in chapter 1 for a fuller description) and these positions change only if an articulator is actively involved in gestural production. So, for /b/ the lips move away from their neutral open position to produce closure, but once the period of gestural activation is over the lips move back towards their neutral position, all things being equal. While the lips are engaged in performing the closure gesture the other articulators remain in their pre-set neutral position (assuming, of course, that they are not themselves involved in gestural production). These neutral settings for each articulator are clearly intended to be to be universally applicable and fully independent of any active gestures which do not involve the articulator in question, but this is, of course, an abstraction from the physical reality. For example, during the target portion of the LIPS gesture for English /b/ the velum is generally tightly closed, thereby preventing any venting of the airflow through the nasal passage which might affect the characteristic release burst (Ohala 1993). Such tight closure is a straightforward enough interpretation of the velum's neutral setting of closure, but it is nevertheless true that this contrasts with the behaviour of the velum during the target portion of a LIPS : **clo** gesture for /v/ where it is in general somewhat lower than for /b/, and lower still for e.g. /w/, yet all three segments have ostensibly the same neutral setting i.e. the velum is in all cases said to be raised. These various non-nasal segments all show different velum activity and appear to have their own values for velum height which can not be accounted for by straightforward mechanical means (Bell-Berti 1980), suggesting that a single absolute neutral setting for each articulator across all environments is somewhat of an idealisation.

Despite the differences in velum height between these segments it is nevertheless true to say that the velum is closed in each case, at least relative to the open velum setting for true nasals, and thus there remains a clear distinction between the role of the velum in e.g. /b/ and /m/. While phonological theories would generally ignore the difference in the behaviour of the velum in /b/ and /v/, any theory of phonetic implementation will need to account for it, and, as Mattingly (1990) has pointed out, AP's claim to be a full model of speech which removes the mapping between phonology and phonetics means that it too must in some way account for these differences in a principled fashion.

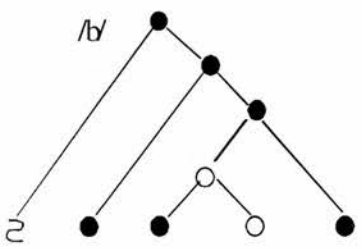
Assuming that /b/ and /v/ each consist of a single LIPS gesture, the differences in the behaviour of the velum in each case must be due directly to the presence of the LIPS gesture. Matters are more complex when more than a single gesture is present, so that while the velum is in a higher position for /b/ than for /v/, it is in a still higher position for /p/ where there is an additional glottal opening gesture. The still higher setting of the velum for /p/ relative to /b/ shows that it cannot be simply the difference in the value for the oral gestures which dictates the neutral setting of the velum but must instead be the combination of whatever gestures are active and the overall state of the vocal tract which results from this combination. Anticipating the discussion below somewhat, let us assume that for both /p/ and /b/ the LIPS : **clo** gesture is the sole head. What the head can then be said to determine is not what the neutral settings for all the articulators are - as /p/ illustrates, it is the combination of all active gestures which determines this[8] - but precisely when these neutral settings occur. For both /b/ and /p/ the neutral settings of the non-active articulators are constrained to be achieved during the target period of the head, in both cases a LIPS : **clo** gesture. In other words, although there are two gestures present in /p/, it is the LIPS gesture which controls the timing of the neutral settings, not the GLO gesture. Thus for English, if a LIPS : **clo** gesture is marked as a head for /b/ there are a set of canonical neutral settings particular to English which the various non-active articulators will take up during the period when the LIPS gesture's target dominates the vocal tract. The same is true of /p/, with the caveat that the presence of the GLO gesture will also affect the neutral settings.

For /b/, then, the LIPS gesture dominates the vocal tract in the sense that any neutral settings of the non-active articulators are constrained so that they are realised

[8]We shall see below that the LIPS gesture for /p/ also controls the behaviour of the GLO gesture.

during the period when the LIPS gesture achieves its target. This is just what we expect given (6) if the LIPS gesture is a head, and we can see the result clearly in the vocal tract hierarchy for /b/ in (7). As we saw in chapter 2, there are a number of problems regarding the interpretation of the VTH in general, but (6) suggests a function for the VTH other than being merely a random selection from an infinite number of possible snapshots of the overall state of the vocal tract at any particular instant. For a segment containing only a single headed LIPS : **clo** gesture, (7) represents the period during which the LIPS gesture dominates the vocal tract. In other words, the structure of /b/ in (5) produces the VTH in (7) where it is true that there is a period during which the LIPS gesture achieves its target and dominates the vocal tract to produce the overall pattern shown. We should note again here, however, that although it is the LIPS gesture in /b/ which demands that the various components of /b/ coordinate specifically to create the structure in (7), given the difference in the values of the neutral articulators in /b/ and /p/, it must be the overall value at the top of the VTH which determines the precise value of the neutral settings of the non-active articulators.

(7)



Similar constraints operate in segments where more than one gesture is present, e.g. /p/, which is distinguished from /b/ by the presence of an additional GLO : **clo** gesture as we have seen, and which therefore might be described as a combination of /b/ + /h/. However, such a description would be misleading as the two gestures in /p/ play quite different roles. The LIPS gesture in /p/, as in /b/, would be best described as providing articulation, but whereas /b/'s phonatory component is provided by the neutral setting of the glottis, the phonatory component of /p/ is instead supplied by the same GLO : **open** gesture which formed /h/. Thus the role of the GLO gesture is here very different from its role in /h/, and this difference in the roles of the oral and glottal gestures underlies the binding theory of Kingston (1990), the assumption being that for a segment such as /p/ it is the oral articulation which determines the behaviour of the glottal opening gesture, and not vice versa. This is further reflected in Ohala's (1990) comments regarding glottalised sonorants and stops, where it is the overall value of the vocal tract at the supralaryngeal level which determines the degree of coordinative freedom which the glottal closing gesture has.

The role of the GLO gesture in /p/ thus seems to be to modify the LIPS gesture. We can see this perhaps more clearly by examining other segments containing the same GLO : **open** gesture but different LIPS gestures. For example /p/ and /f/ both contain LIPS gestures, a closed one in the case of /p/ and a critical one in the case of /f/, and each is accompanied by an overlapping GLO : **open** gesture which provides a period of voiceless phonation throughout the period during which the LIPS gestures' targets are achieved. Nevertheless, the nature of the GLO gesture is not identical in /p/ and /f/. For example, the glottal opening gesture of voiceless nonaspirate stops such as /p/ is often timed to reach its peak point of opening just before the release of the closure gesture, perhaps resulting in a brief period of postaspiration (Catford 1977). In contrast, for voiceless fricatives the glottal opening gesture instead reaches peak opening generally around the midpoint of the oral critical gesture, thereby providing as large a degree of airflow as possible throughout the period of oral constriction, and in addition the degree of glottal opening is usually significantly larger than that found for voiceless stops. Despite these differences it is true to say that we have essentially the same GLO gesture for both /p/ and /f/, in precisely the same way as the 'same' neutral setting of the velum is found for /p/ and /w/.

In both /p/ and /f/, then, a LIPS gesture influences the degree of opening, duration and so on of an accompanying GLO gesture, and in addition controls the manner in which the two coordinate. In other words, during the period when the LIPS gestures are active for /p/ and /f/ they dominate the vocal tract just as the LIPS gesture for /b/ dominates the vocal tract, though in addition to controlling the behaviour of any neutral settings they also constrain the GLO gestures so that their targets are achieved during the time that the LIPS gestures' targets are achieved. This unequal relationship between the oral and glottal gestures suggests the representations in (8), where the LIPS gestures are marked as heads, dominating the non-headed GLO gestures. Thus, the LIPS gestures may be said to govern the GLO gestures, or alternatively the GLO gestures may be said to be governed by (or be dependent on, or modify) the LIPS gestures.

(8)

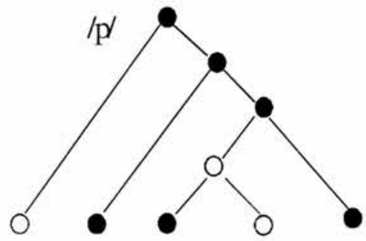
a) /p/ - LIPS^{H} : **clo**, GLO : **open**

b) /f/ - LIPS^{H} : **cri**, GLO : **open**

The relationship between the LIPS and GLO gestures in /p/ is now one of head and dependant, and it this phonological relationship which constrains the physical relationship between them. The role of non-heads is to modify heads during the period when heads dominate the vocal tract, and therefore non-heads do not themselves dominate the vocal tract at any point during their activation. Thus the structure in (8a), with its single headed gesture, generates a single period of domination of the vocal tract and hence a single VTH as in (9). This VTH differs from that of /f/ only in that closure rather than a critical setting is percolated up from the LIPS gesture, and neither (9) nor the principles in (6) can be said by themselves to detail exactly the coordination between the LIPS and GLO gestures. What we can infer from (6) and (8a) is that the targets of non-heads are achieved wholly during the target period of heads; in other words, non-heads are completely overlapped by heads. For example, for /p/ where there is only a single head dominating a single non-head, the non-head i.e. GLO : **open** must be completely overlapped by the head i.e. LIPS : **clo**, as if this were not the case the GLO gesture would extend beyond the LIPS gesture and thereby generate its own VTH, in which case it would be a head.[9]

[9]This is somewhat of an idealisation, referring as it does to a segment in isolation, as within speech there may very well be other considerations which affect the precise coordination of all gestures.

(9)



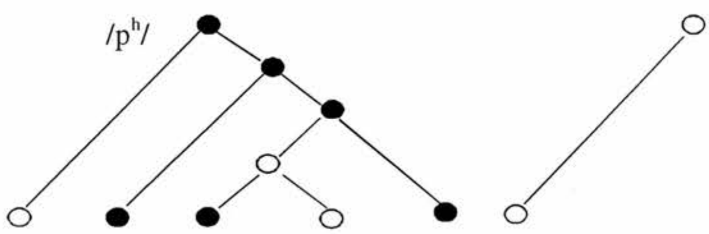
The structures in (8), together with the principles in (6), are general constraints on the coordination of headed and non-headed gestures, but are not in themselves sufficient to precisely determine coordination between the LIPS and GLO gestures of /p/ and /f/. Taking /p/ as an example, the structure of /p/ in (8a), where LIPS : **clo** is a head, means that there is a period when it alone dominates the vocal tract, and that during this period the LIPS gesture controls the behaviour of all other gestures, i.e. it specifies whether any other gestures present in the segment coordinate with it and, if they do, what that coordination is. The GLO gesture, on the other hand, is not a head and, as we have seen, this means that the GLO gesture is effectively completely overlapped by the LIPS gesture. What is missing is any information specific to a headed LIPS : **clo** gesture as opposed to any other type of head.

As noted above, there are differences between /p/ and /f/ both in the coordination of glottal opening gestures with the respective labial gestures, and with the glottal opening gestures themselves, and it is the reasons behind such differences that e.g. Kingston's binding theory is designed to answer. The headed structures proposed here provide some general constraints e.g. LIPS and GLO in both /p/ and /f/ must show complete overlap, but the differences between /p/ and /f/ are more specific. The differences between the coordination of gestures in /p/ and /f/ must be part of the overall specifications of AP if it is to be a full description of speech, just as the differences in the exact values of the neutral settings must also be included, and the same is true of any full phonetic implementation. Given this assumption, and following (6bii), part of the information that each language must contain is such that for every gesture X which is a head, if gesture Y occurs within that same segment then gesture X will coordinate with it or not; if X does coordinate with Y, then that coordination will be of the type Z.

For example, perhaps universally we can claim that throughout the period during which a LIPS : **clo** gesture dominates, any GLO : **open** gesture must also achieve its target, and in fact we can generalise this to the domination period of any oral closure gesture (and possibly any oral critical gesture), with the result that for any segment containing a headed oral closure gesture and a glottal opening gesture, the period during which the oral gesture is a head will be accompanied by voicelessness. This is the kind of control of other gestures implied by the definition of a head in (6), and for every gesture which can be a head in a language there are similar constraints, some language specific, others universal. For /p/ this means that the domination period of the LIPS gesture would be characterised by voicelessness regardless of the fact that the two must completely overlap in any case, given the single headed structure. It is important to note, however, that the control of a glottal opening gesture by an oral closure gesture only demands that the glottis be open during the period when the oral gesture is a head. It does not demand that the glottal opening begin or end at any particular instant relevant to the oral closure, only that it be open while the closure dominates the vocal tract. Thus, glottal opening could be fully achieved before oral closure has even begun, and the glottis could remain open long after oral closure has been released. What further constrains the glottal opening gesture is, as we have already seen, the fact that for /p/ the sole head is the LIPS gesture and thus the GLO gesture must be completely overlapped by the oral closure. Ultimately these various constraints are what produce the required output.

In contrast to /p/, whose head-dependant structure produces a single VTH, as we saw in chapter 2 a description of /p^h/ in the same terms clearly requires two separate hierarchies as in (10), containing the same period of closure which is characteristic of /p/ but in addition a following period of what can be described either as voicelessness or as noise, but is better described as both. The gestures for /p/, which show complete overlap, constitute a single segment in the same way as /p^h/ where the gestures show only partial overlap, and as for the other segments we have described so far, the physical relationship between the gestures must be a direct consequence of the phonological relationship between them. What we require, then, is that the phonological structure of $/p^h/$ will demand that the two gestures be offset without having to rely on the arbitrary specification of partial overlap.

(10)



Looking back to the description of unaspirated and aspirated stops in chapter 2, there they were said to show complete and partial overlap respectively, where partial overlap is derived by starting with complete overlap and then gradually sliding the gestures apart. Following this reasoning to the letter would suggest that for $/p^h/$ the GLO gesture would begin at a later point in the cycle of the LIPS gesture than it would for /p/. However, this is not the case. Instead, the glottis begins to open at roughly the same time relative to the LIPS gestures for both /p/ and $/p^h/$, the difference between the two lying in the maintenance of glottal opening beyond the release of the oral closure for the aspirated stop. In addition to the many problems with the use of the complete, partial and minimal overlap system, this seems quite incompatible with the definition of partial overlap as it stands at present.

Viewed instead from the position of headed structures, one solution immediately suggests itself, that is that as $/p^h/$ appears to be composed of two consecutive events, then it ought also to contain two heads as in (11). How, though, are we to interpret such a structure? In line with the structures discussed above we can assume, as both gestures are heads, that there is a period during which the LIPS gesture alone dominates and a period during which the GLO gesture alone dominates, as described in (6). The only way in which this can be achieved is if the periods during which the respective gestures dominate are consecutive, i.e. one gesture dominates and then the other gesture dominates, but not both simultaneously.

(11)

/p^h/ - LIPS^H : **clo**, GLO^H : **open**

Recall that for /p/ the presence of only two gestures and a single head means that the two gestures must show complete overlap (or rather that the GLO gesture must be completely overlapped by the LIPS gesture). Once there is no longer a single head but two heads this requirement no longer holds, so that in theory the LIPS and GLO gestures in (11) need only minimally overlap each other in order to be compatible with a single segment analysis. What determines the actual coordination are the principles outlined above, i.e. for every gesture which can be a head we must provide information as to how it coordinates with any other gestures within the same gesture. This information, as already noted, must be included in any complete model of speech, including AP. One particular part of this information noted above was that for any headed oral closure gesture, if there is also a glottal opening gesture within the same segment then the period during which the oral gesture dominates will be overlapped by the glottal gesture. Therefore, during the period in which the LIPS gesture dominates, the oral closure of /p^h/ will be characterised by voicelessness in the same way as in /p/, i.e. oral obstruent gestures treat any glottal gestures present within the same segment as if they were phonation. In fact, the oral closure portions of /p/ and /p^h/ are to all intents and purposes identical, as we have seen.

No such assumption can be made, however, for the period during which the GLO gesture dominates. In other words, when a GLO **clo** gesture dominates it does not overlap any other gestures, with the result that for /p^h/ the period during which the GLO gesture dominates will not be characterised by the presence of the LIPS gesture, and this claim seems to be universally true.[10] Headed glottal gestures do not assume that any oral gestures within the segment are periods of articulation which must cooccur with them, and the result of this for the structure in (11) is the creation of a period of voiceless labial closure and a separate period of voicelessness, these two periods ordered with respect to each other.

We now have a structure which can be interpreted as either /p^h/ or /^{h}p/, that is as either a postaspirated or a preaspirated stop, and there seems to be no immediate reason to prefer one over the other. Why, then, is /p^h/ almost universally preferred? As Kingston (1989) suggests, the answer to this almost certainly lies in the nature of plosives, and more particularly in the role of their release bursts and the role of the glottis in enriching the cues to the identity of the stop at its release. However, the fact remains that there are languages in which the structure in (11) can be realised as a preaspirated, rather than postaspirated, stop. As Kingston & Cohen (1992) point out the demand for explicit coordination leaves AP apparently unable to rule out phonological contrasts between pre- and postaspirated stops, while the lack of such a strict physical basis in feature geometry, coupled with the (apparent) absence of ordering in its representation of aspirates, means that it predicts that no such contrast should be possible. If, though, we adopt the structure in (11) we see that AP can make the same general claim as FG, that is that languages cannot contain a phonological contrast between preaspirated and unaspirated stops. Both AP and FG (at least at the level of phonetic interpretation) must simply specify that in any particular environment the oral and glottal parts of an aspirated stop are consecutively realised in the appropriate fashion.

The AP representation does appear to have some advantages over that of FG, in both its avoidance of an equivalent of [+spread glottis] and in its insistence that there be two events and therefore that there be ordering of the gestures, but without carrying any implications as to what this ordering should be. The fact that phonetically the glottal and oral gestures are in fact ordered has to be ignored by

[10]We shall see in chapter 5 some cases in which headed glottal gestures and oral gestures are

feature geometry as it lies outside its domain (and as Löfqvist & Yoshioka (1981) point out, the use of [+spread glottis] is both phonetically inaccurate and introduces unnecessary complications), while AP rules out the contrast between pre- and postaspirated stops but at the same time captures the fact that the gestures must be offset. Both pre- and postaspirated stops then have precisely the same gestures and the same head-dependant relationships, the difference between them lying elsewhere in the phonology (see chapter 4 for further discussion).

We thus have a set of headed representations for simple segments which account for the general physical coordination of the gestures within segments. In addition to this, however, AP must include a great deal of additional information to account for the behaviour of each individual headed gesture within a language. Some of this information we have already seen, e.g. headed oral closure gestures are voiceless if a glottal opening gesture occurs within the same segment, and it should be a relatively simple task to identify many more such general principles. What is more difficult is to identify the precise physical patterns which differ from language to language. In the remainder of this chapter I will attempt to shed some light on both of these areas for more complex segment types.

3.2.2
Complex Segments

Segments such as /p/ and /pʰ/ involve gestures within separate subsystems, but segments may also consist of multiple gestures within a single subsystem.[11] We saw in chapter 2 that the representations of complex segments in both Feature Geometry and AP were inadequate in a number of ways, each making the claim that complex and contour segments were special classes to be distinguished from other

---

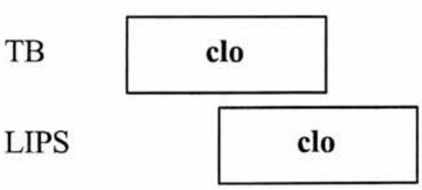
coextensive. This does not affect the generality of the claims made here, however.

[11] It may be worth recalling here the distinction between complex and contour segments. Complex segments such as /kp/ are said to exhibit no phonological ordering between the place features (here [dorsal] and [labial], while they do show phonetic ordering between the two. In contrast, contour segments such as /nd/ contain opposite values for a single feature, in this case [+nasal] and [-nasal]. As these two features can not be simultaneously realised they must be phonologically as well as phonetically ordered.

segment types but each failing to make that distinction clear. However, there is again no a priori reason why this distinction between simple, complex and contour segments should be made in AP, but if this distinction is removed we must determine what consequences it would have for our model.

Maddieson & Ladefoged (1989) show that for the labiovelar stop /gb/, ordering, at the phonetic level at least, is essential and universal. (12) shows in gestural terms their interpretation of the ordering of the two gestures, where the velar closure begins slightly before the labial closure and ends slightly before the labial closure is released. Although the period of time between the release of the two gestures is relatively short - typically one frame when viewed in terms of x-ray cine film - it is clear that there is no difference between the pattern seen in (12) and that seen for postaspirated stops, i.e. two gestures partially overlapping to give two separate events. The nature of the gestures leads to quite distinctive audio-acoustic events in the two cases - a combination of gestures in two separate tubes as compared to two gestures in a single tube - but they are identical in terms of the number of events the gestures produce. The difference between them lies in the fact that this is the only type of segment that two oral closure gestures can create, there being no distinctions comparable to those possible using different subsystems such as voicelessness versus aspiration + voicelessness, i.e. there are no ordering distinctions. The explanation for this is quite clearly not to be found in the articulation, there being no physical reason why the two gestures should not begin and end simultaneously, but instead we need to look at the problem from the listener's perspective.

(12)



The formants of the vowel preceding /gb/ are primarily affected by the velar closure, as it is this which overlaps the vowel to the greatest extent, while it is the labial gesture which is the greatest influence on the formants of the following vowel. This is as we would expect from the gestural ordering. The dominance as it were of complex segments by labiovelars is due to a number of factors. The most important of these are that the use of two separate articulators, LIPS and TB - complex segments using TT and TB are rare - gives enough flexibility to make coordination possible, and more importantly the acoustic effects of labial and velar gestures are similar and each reinforces the other, making the segment more distinctive. Were we to have TT and LIPS closure gestures for e.g. /pt/ which were overlapped to the same extent as the gestures for /gb/, what we would in fact find is /p/ as the labial gesture, even at much lesser degrees of overlap, would acoustically dominate (Byrd 1992). The vowel formant transitions into a TT consonant are much smaller than those into a LIPS consonant and would be dominated by it, and the relatively much greater speed of the TT movement would easily lead to it being completely hidden by the slower labial closure, as Byrd shows. This does not rule out the existence of such segments, implying only that the fine coordination necessary would make it difficult to maintain their distinctiveness.

We have no such problems for complex labiovelar stops, as we can have almost complete overlap and still maintain their distinctiveness, TB and LIPS gestures being relatively robust in comparison to TT gestures. However, as we can see from Maddieson & Ladefoged's data, while the mutual reinforcement effects of labial and velar gestures might lead to labiovelar gestures being favoured as complex segments, the effects on the surrounding vowels seem to be little different from the effects of a sequence of /g + b/. The main difference between a /g + b/ coda + onset sequence and a complex cluster /gb/ (discounting any 'phonological' effects such as their respective behaviour during prenasalisation, and the greater amount of overlap found in complex segments) is that in a coda-onset sequence both closures have an audible release while in a complex segment we have apparently only a single release. In other words the release of one gesture is hidden by the other. Obviously the release of the hidden gesture is only important in the sense that when the second gesture is released the other gesture has already been released. They must not be released simultaneously. If it were the mutual reinforcement that was vital we would expect both closure gestures to be achieved and released simultaneously, yet this is never the case and AP provides us with an answer as to why this should be so.

Despite the robustness of labial and velar closure gestures, if we were to have complete rather than partial overlap of the two we would lose any distinctiveness gained by coordinating the two gestures. One or the other of the two gestures would dominate, with the result that it would be indistinguishable from a single /g/ or /b/. Rather than viewing the ordering of the gestures in a complex segment as a secondary phonetic phenomenon, it is clear that we should see it as primary. We find complex segments because they are in fact relatively easy to make, but in order to 'survive' as simple segments rather than clusters their constituent gestures must be offset, which allows us to interpret both the TB and the LIPS gestures of /gb/ as heads. Any phonological unordering (if such is present) is simply a result of the fact that it is a single segment. The representation in (13) reflects the fact that /gb/ is both a single segment and that the two gestures produce, as it were, two separate events. There is thus no distinction to be made in terms of internal structure between aspirated stops and complex segments, and between these and plain unaspirated stops such as /p/. They are all single segments, their slightly different natures being a result of the number of headed gestures each contains.

(13)

/gb/ $\quad$ TB^{H} : **clo**, LIPS^{H} : **clo**

This analysis is not overtly different from that of feature geometry seen in chapter 2, yet in fact it differs from feature geometry in a number of significant ways. In AP we have at present five separate articulators, though undoubtedly more are needed (Browman & Goldstein 1992a), and these articulators are further placed in three different subsystems, namely glottal, velic and oral. Gestures from two separate subsystems can be coordinated to form segments, and indeed the combination of oral and glottal gestures is universal, yet at the same time there is no reason whatsoever that we should not expect the different articulators of the oral tube to coordinate in the same way, which of course they do. So far we have seen that two oral gestures of the same constriction degree can form a segment, but the nature of the gestures makes it difficult to keep them acoustically distinct from other simple segments, making them relatively rare in natural languages. From an articulatory point of view there is no reason why we should not find such segments; other considerations account for their rarity.

The gestures involved in a complex segment such as /gb/ are identical in all respects to those in the sequence /g + b/, reflecting the relative independence of the different articulators, and therefore we would expect to find segments in which the constriction degrees of the constituent gestures differ, such as /ps/ of Margi or Kabardian (Padgett 1991). For standard Feature Geometry this poses a number of problems as we have seen, as [cont] and [place] are at different levels in the tree, and as a result it is difficult to percolate the correct values of [cont] to the appropriate [place] node. No such problems exist for AP. The two gestures in (14) for /ps/ do not in any way share different values for a single branching [cont] node; they are separate and independent gestures which combine to form a single segment but which do not rely the one on the other for any of their values. (14) produces the structure in (15) and is no different in its basic structure from /pt/. Rather than differing in having an additional [cont] node, they differ in having different gestures, and no ad hoc pointer system is necessary, thus the difference between complex segments with the same value for [cont] and coarticulated segments with differing values for [cont] is done away with.[12] Padgett (1991) adopts a similar approach in terms of feature geometry.

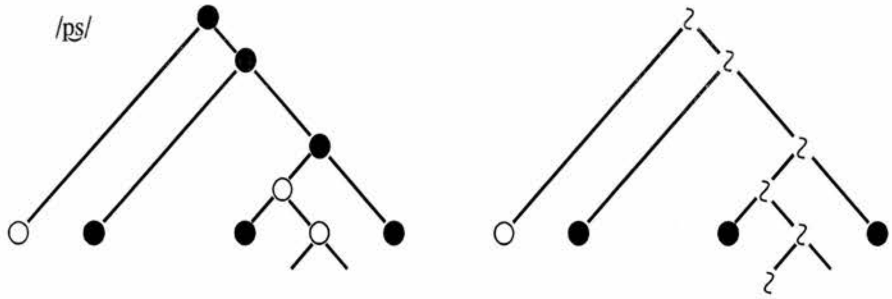
(14)

---

[12]Both gestures here are heads, unlike Sagey (1986) where one or the other but not both can be a head. This does not preclude one gesture implying another, if underspecification were desirable. Padgett (1991) makes the same point in his discussion of Kabardian.

/ps/ LIPS^{H} : **clo**, TT^{H} : **cri**

(15)



In summary, AP claims that the complex segments and similar sequences of consonants differ in the phonological relationships between the gestures, the phonological relationships resulting in the difference in coordination and thus in the degree of overlap of the respective gestures. By increasing the overlap of labial and velar closure gestures we change from /g + b/ to a segment /gb/, and presumably this is how many if not all complex segments are created (Laver 1994). The increase in overlap prevents the release of one of the gestures from being audible, distinguishing in this way between sequence and segment. Further, complex segments must also of course have both gestures belonging to the same subsystem, but these multiple gestures behave as if they formed as it were a complex value for Place rather than being ordered values. This is reflected in their behaviour during prenasalisation, /gb/ typically becoming /mŋgb/ (though see section 3.4). It is the oral tube which is a head, and the gestures in that tube are themselves heads with respect to their coordination with each other, but there is no sense in which the two gestures branch as in feature geometry so we expect assimilation processes to involve both gestures equally (nasal place assimilation is discussed in greater detail in the following section).

The advantage of this kind of representation is obvious. Both heads in /gb/ are in the oral subsystem and behave as a unit, yet at the same time the nature of the headed structure of AP demands that the two gestures be offset. Perhaps the primary requirement of so-called complex segments is that the gestures (or features) must not be articulated simultaneously but must be offset. If they are not then they cannot be distinguished from segments involving only a single oral gesture, so that /gb/ will be confused with /g/ or /b/, and /ps/ with /p/ (the pressure build up caused by the labial blockage would prevent the airflow necessary for frication). AP directly reflects this fact without preventing the 'phonological' unordering necessary, whereas feature geometry cannot, making AP simpler and more explanatory and thus preferable as both a description and a theory.

3.2.3
Secondary Articulation

Segments of the type discussed here are compatible with the strong view of the oral gesture ventured above, i.e. that gestures in the oral tube are always heads, but there is one class of segments which forces us to reevaluate this position. Up till now, the complex segments we have been discussing have involved the coordination solely of consonantal gestures, but far commoner are those which coordinate consonantal gestures with vocalic gestures. Superficially these may give the appearance of showing the same kind of ordering that is present for complex segments. Many languages realise secondary palatalisation by releasing the palatal constriction relatively slowly, and the palatalised consonant's offset phase is characterised by a '[j]-like offset' (Laver 1994), or if viewed from the following vowel's perspective, the vowel receives a [j]-like onset. This is typically the case in the palatalised stops of both Russian and Irish, for example, and both also tend to produce velar approximant off-glides from their series of velarised consonants. At face value it seems that we have a situation identical to that of other complex segments and that therefore we should treat these in the same way by making all oral gestures heads.

Unfortunately the presence of these off-glides is by no means universal, and even within languages there is variation. The velarised consonants of Russian, while often producing off-glides, may instead have the secondary vowel articulation completely overlapped by the accompanying consonantal gesture, so that a velarised

/tˠ/ can be realised with or without an off-glide. Further, rather than producing off-glides many consonants seem to always show complete overlap. This is commonly the case with e.g. palatalised /kʲ/ which in many languages, including again Russian and Irish, may be realised as [c], or alternatively the constriction location may instead stretch from the velum to the hard palate. In neither case is there any off-glide, yet the same dialect or even the same speaker may show all of these patterns. This suggests that unlike complex segments involving solely consonantal gestures, for segments such as /kʲ/ it is not important whether or not two separate events are created, rather the simple coordination of the two gestures is all that is required.

If we look at labialised consonants, which involve two physiologically independent articulators, the labialisation tends to spread to the surrounding segments, showing a particular propensity for anticipatory spread (Catford 1977). The coordination between the labialisation and its accompanying consonantal gesture is relatively lax, allowing the two to drift apart to a large extent without causing confusion with other segments. There is no language which contrasts labialised segments with the sole contrast being whether or not the labial gesture shows partial or complete overlap with the consonantal gesture, just as there is no language which phonologically contrasts pre- and postaspiration. So-called secondary articulations of this kind chiefly signal their distinctiveness from otherwise identical segments without secondary articulation by the vowel-to-vowel formant changes generated, and ordering of the respective gestures would not significantly affect these formant changes, at least not enough to be contrastive. So while two coordinated consonantal gestures must be offset and therefore partially overlap in order to form distinctive segments, coordination of consonantal with vocalic gestures places no such demands, requiring only that the gestures overlap.

In fact, it seems reasonable to assume that for such segments it is the consonantal gesture (or gestures) alone which is the head, the vocalic gesture being dependent upon it.[13] As the term secondary articulation suggests, for /kʲ/ the

[13] Compare the similar conclusions of Dependency Phonology (Anderson & Ewen 1989).

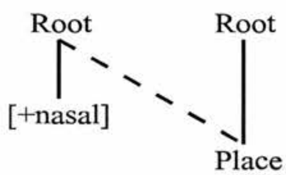
consonantal TB closure gesture is the primary articulation, the palatalisation acting to modify it in the same way as the glottal opening gesture modifies a simple /k/. We should therefore reflect this in the phonological relationship between the gestures. The presence of offglides is due to a number of factors, including the inherently longer duration and slow release of vocalic gestures, with the possibility that it may be also due, at least in part, to the pervasive influence of vowel-to-vowel articulation. We recall that vowels provide a continuous bed upon which consonants lie, temporarily obscuring them, with vowels being coordinated first of all with each other and only secondly with the surrounding consonants. If we analyse secondary articulations as involving vocalic gestures as well as consonantal then we might assume that the vocalic gesture of the consonant coordinates in one way with the following vowel, and the consonantal gesture coordinates with this vowel in a slightly different way. The fact that one of the vocalic gestures is part of a consonantal segment is irrelevant. Any offglides which may be created are, as it were, the result of purely mechanical changes in the position of the tongue or lips as they move to the positions required for the following vowel (cf. the description given of a dialect of Irish by Mhac an Fhailigh 1980).

The question of the representation of secondary articulation has been widely discussed in feature geometry (see Ní Chiosáin (1994) and references therein), especially as to whether it should be represented with a separate set of articulators than those used for consonants, and if so how these separate articulators coordinate with one another. The position suggested here is that secondary articulations are effectively vocalic gestures coordinated with consonantal gestures and as such they coordinate differently with the surrounding vocalic gestures which function as vowels. Precisely how this is achieved I will not attempt to answer here, as we still need a greater understanding of how vowels coordinate in general, and in particular we must provide an answer as to why such secondary[14] vocalic gestures appear never to be heads, but I believe that secondary articulations do not raise any serious challenge to the theory of heads outlined here.

3.3
The Representation of Nasals
3.3.1
Simple Nasals

The proposals outlined above represent one possible solution to the problem of integrating both categorical and gradient information into a single coherent system, but so far there is no evidence that such a formulation can be extended beyond a mere description of segments into the realm of phonological rules proper. This question of AP's ability to model phonological rules was raised by Clements (1992). Given that the VTH is constructed solely on the basis of anatomical considerations, Clements claimed that it lacks the necessary internal structure that we find in FG. Of particular concern to Clements is the lack of a node akin to the [place] node of FG, in particular its function of grouping together the various nodes dependent on it, and its abstract nature which allows it to take part in assimilation processes. For example, in Yoruba and many other Bantu languages we find pervasive nasal assimilation and prenasalisation, so that e.g. N + bɛ gives m̩-bɛ,[15] and FG represents this simply as in (16), where the use of abstract nodes allows the nasal's Root node to link directly to the [place] node dependent on the following Root.

(16)



In contrast, the VTH of AP is based solely on the anatomy of the vocal tract and phonological patterns must arise from these physical considerations, and there is

---

[14] We must also, of course, agree upon a definition of what the term 'secondary' means in this context.

no direct equivalent of the [place] node available to either spread or to act as a target for spreading. The hierarchy of FG, while matching that of AP in many ways, is not bound by physical considerations in the same way, as we have seen. The question then arises regarding prenasalisation in Yoruba as to how precisely, in Clements' own words, 'all the oral gestures of any consonant (to the exclusion of its laryngeal and velic gestures) must overlap a preceding nasal' (pp 186-7 op. cit.).

In order to provide an answer to this, we must first be able to answer the question 'what is a nasal?'. This should be a relatively straightforward matter, given the structures proposed in this chapter. If we take a segment such as /n/ we can see its apparent similarities to other segments we have already discussed. Most obviously it seems to consist of a single event in the same way as does /t/ or /l/, as opposed to the two events of a segment such as /gb/, and we should be able to reflect this fact in our representation. For such simple nasals we have two gestures, one lowering the velum, the other forming complete closure in the oral tube, suggesting, for example, that /n/ will have the structure VEL^{A} : **open**, TT^{H}: **clo**. This assumes again that it is the oral gesture which is the head, the velic gesture being dependent upon it and thus creating a single event. This interpretation raises a number of interesting questions.

First of all, it is nasal segments' phonological behaviour that allows us to think of them as single segments, yet the phonetic facts show that they often behave in ways very different to those of other single segments. Vaissière (1988) shows that for a (word-initial) sequence CVN such as 'ban' in (American) English, the velum reaches its maximum degree of lowering during the vowel and is often already beginning to rise as oral closure is achieved.[16] This would perhaps give us reason to think that in fact we have here two events rather than one, which would result in problems distinguishing between this and the more 'traditional' type of nasal contour,

---

[15]The nasal may or may not be syllabic, but there is no difference in the structure in either case.

[16] Of course, when the nasal is followed by another consonant e.g. 'bandit' many American accents have no oral closure whatsoever for the nasal, having instead a nasalised vowel.

especially in a system such as AP where phonetic ordering must be shown. In fact, the problem disappears when we compare the velum with the other articulators.

The independence of the velum from the oral and glottal articulators, combined with its (lack of) speed of raising and lowering, both demand and allow a great deal of variation in its timing in relation to the gestures around it. The relatively slow velocity of velum lowering means that lowering must often begin some time before the relatively fast closure gestures which accompany it, otherwise oral closure could be made and released before the velum has had time to lower sufficiently to create distinctive nasalisation, and similarly the slow velocity of velum raising means that raising can begin before oral closure is achieved and still remain low enough to maintain nasality throughout the closure and even beyond. In some respects it mirrors the behaviour of vocalic gestures in secondary articulations in that while the active gestures might not be completely coextensive, there is no ordering involved. As long as the velum remains distinctively low during oral closure it is irrelevant precisely when lowering and raising begin and end, so there is no ordering in the sense developed here (see Laver 1994 for a similar conception). AP reflects these facts by specifying that the gestures form both a discrete unit and a single event, and the velum is thus constrained to remain lowered throughout the period of oral closure, but at the same time the phonology contains information about the language's (and the individual speaker's) stiffness settings for the velum, its height in the surrounding segments and so on, all of which are necessary to compute the velum's exact pattern of movement and are reflected in the velum's relative freedom of movement. This is what is implied in the marking of the oral gesture as a head. In this way then AP reflects both the categorical nature of simple nasals and at the same time the gradient nature of the velum's coordination with the surrounding gestures.

As far as the identity of the head for nasal segments is concerned, the evidence seems to again show that it is the oral gesture which is primary for all nasals. Although the nasal passage forms a clearly distinct tube and the velum itself shows a large degree of articulatory independence, in many ways it is dependent upon the presence of other articulators. The glottal tube, like the velic, is placed above the oral in the VTH, and, as we have seen, gestures formed within it can form segments such as /ʔ/ and /h/, but there are no such segments formed solely of velic gestures. The apparent existence of floating nasal segments (comparable to floating glottal segments like hâche aspirée in French) such as in Guaraní (Piggot 1992) might suggest that velic gestures can also occur alone, but in fact such segments show that they are dependent on oral gestures for their realisation. Prenasalisation can also be seen in terms of a floating velic gesture, and if a such a segment does not receive an accompanying oral gesture it typically receives a default gesture (default place in FG) or deletes altogether (Padgett 1994). Therefore it seems as if velic gestures are in a sense subordinate to oral gestures in that they must cooccur with them. The reverse of course is not true, providing further evidence for the view that oral gestures for consonants are always heads.

Simple nasals can thus be seen as being the same type of structure as the other segments we have seen so far, and I suggest that nasal contour segments should bear the same type of relationship to simple nasals as /g/ and /b/ do to /gb/, despite claims that contour segments and complex segments should be distinguished by phonological versus phonetic ordering. Simple nasals can be viewed as being stops with an additional side chamber which prevents the build up of pressure which would otherwise lead to a burst at release. For /n/ the lowered velum acts to modify an oral closure gesture in the same way as the glottal opening gesture for /k/ and the secondary articulation for /k^j/. This leads us to represent /n/ as TT^H : **clo**, VEL : **open**.

3.3.2
Prenasals

Given that for simple nasals there is a single event and therefore a single head, for prenasals, where there are two consecutive events, we would expect there to be two heads. In this situation, the oral and velic gestures need no longer be completely coextensive, but given that they still form a single segment there will inevitably be some overlap. In stops such as /p^h/ the result of the glottis being a head is to create a separate event characterised by the sole presence of the glottal gesture, a fact made possible by the ability of glottal gestures to exist independently of gestures in any other tube. How, though, should we account for the coordination in segments such as /m͜b/?

Given that /m͜b/ clearly contains two consecutive events in the same way as do /p^h/ and /gb/, we would expect it to have the structure in (17). This structure, following the principles of (6) above, results in a period during which the LIPS gesture dominates together with a period during which the VEL gesture dominates. Again, just like /p^h/ the precise ordering between the two events must be determined language specifically as while both prenasals and postnasals are possible, no language contrasts between the two. In addition, following the arguments above regarding the relationship between oral and nasal gestures, we would expect any event during which a velic gesture was active to be accompanied by oral closure universally. Obversely, while for /m/ the oral head is completely overlapped by the velic gesture, this does not imply that the same be true when the velic lowering gesture is also a head. In fact, we can make the claim that universally the domination period of headed oral closure gestures is never characterised by overlapping with a headed velic lowering gesture. The prenasalisation of /b/ to /m͜b/ now simply involves the addition of a headed velic gesture, so that /b/ LIPS^H : **clo** becomes /m͜b/ LIPS^H : **clo**, VEL^H : **open**.

(17)

LIPS^H : **clo**, VEL^H : **open**

It is worthwhile reiterating here that there is nothing in the nature of the principles of (6) which rules out the possibility of the domination period of the LIPS gesture in (17). What (6) and (17) demand is that there be two separate and consecutive events, one dominated by the LIPS gesture, the other by the VEL gesture. The nature of these two periods, in particular the realisation of the neutral settings of non-active articulators and the coordination (or its absence) between the active gestures, must be determined separately from empirical observation. While there seems to be much evidence, for example, of the dependence of velic gestures on oral gestures, the precise reasons which underlie this dependence remain opaque. Similarly, compare the wide range of evidence cited by Padgett (1991) as to the existence of the universal constraint *[+nas, +cont] ruling out nasal fricatives. Very few apparent exceptions to this constraint exist, yet there seems little in the way of either physical restrictions or the nature of distinctive features which would account for the constraint. This is not to say that the reasons for such constraints cannot be found, just that they are not always immediately determinable, and the same is true of the constraints proposed so far for AP. What we can do here is to identify as many constraints in AP as possible, providing explanations for them where we can.

Turning to more complex prenasals, Scottish Gaelic, as noted earlier, has both aspirated and unaspirated stops, and both sets can undergo prenasalisation (also known as eclipsis) as in (18). For a voiced stop /b/ to become /mb/ as in (18), the suggestion is that we simply add a headed VEL gesture, so we expect the same principles to apply with the voiceless stops of Gaelic in (18). /p^h/ as we have seen consists of two gestures, $LIPS^H$: **clo** and GLO^H : **open**, creating two events, the glottis being open throughout the period of closure; the same gestures in /p/ are effectively coextensive with respect to each other, creating a single event. When prenasalised, therefore, we might initially expect the same principles of coordination to apply for N + /p/ as for /p/, with the GLO gesture coextensive with the closure gesture, but as we can see in (18) the period during which the velic gesture is active is voiced and not voiceless as might be expected, and in fact the nasal part of prenasal stops seems to be almost universally voiced (Herbert 1986). This is initially surprising given that the AP representation for /mp/ is $LIPS^H$: **clo**, VEL^H : **open**, GLO : **open**, with a glottal opening gesture which would suggest that the result of prenasalising /p/ would be [m̥p], with the GLO gesture overlapping the LIPS gesture, which is itself overlapping the VEL gesture, rather than [mp]. This is precisely the point made by Clements in that it seems impossible to assimilate manner and place in AP without also assimilating the laryngeal features.

(18)

a.

ball	[pauɫ] n. 'ball'	am ball	[^{m}pauɫ] 'the ball'
dall	[tauɫ] n. 'blind person'	an dall	[^{n}tauɫ] 'the blind person'
gall	[kauɫ] n. 'foreigner'	an gall	[$^{\text{ŋ}}$kauɫ] 'the foreigner'

b.

palla	[p^{h}aɫa] n. 'shelf in rock'	am palla	[^{m}p^{h}aɫa] 'the shelf'
toll	[t^{h}ouɫ] n. 'hole'	an toll	[^{n}t^{h}ouɫ] 'the hole'
call	[k^{h}auɫ] n. 'loss'	an call	[$^{\text{ŋ}}$k^{h}auɫ] 'the loss'

FG avoids this problem altogether by using the [place] node as in (16) above. Assimilation is to this node only, there being no need for the nasal to assimilate to the other nodes such as [cont] or [laryngeal], and as a result the nasal is left free to receive a [laryngeal] node of its own, which will generally be [+voice]. As Clements points out, this kind of analysis seems possible only if we have a separate [place] node. The solution for AP lies, I suggest, in what makes us expect the GLO gesture of /p/ to overlap the nasal portion of a prenasalised stop.

It is the single-headed nature of /p/ and other simple voiceless stops which accounts for the fact that their oral and glottal gestures are coextensive. As the GLO gesture is not a head it cannot create a separate event and must be active (largely) during the period in which its head gesture is active, but the exact coordination between the two gestures of /p/ is not detailed by this headed structure in any more precise way, the headed structure only demanding that they coordinate in such a way as to form a single distinctive event.[17] As we saw earlier for the aspirated stops of Icelandic and Hindi, the precise manner in which languages create these events is

[17]The glottis has a similar, if more restricted, freedom to the velum, in that complete overlap is not always necessary. English voiceless stops, while part of a series distinguished by presence or absence of voicing, are nevertheless usually, though not always, aspirated, indicating that exact coordination between the glottal and oral gestures is unimportant. This kind of aspiration does not require the presence of a headed glottal gesture.

specific to the language and must be specified separately, though cross linguistically we will see that certain patterns dominate for various reasons (Kingston 1990).

Goldstein (1990) suggests that for English the onset can contain only a single ballistic opening of the glottis, i.e. one glottal opening gesture, so that /s/ + stop clusters have to share one GLO gesture.[18] The result is a compromise between the two different types of coordination, the glottis reaching peak opening at the boundary of the two segments, later than the midpoint of the fricative and earlier than is normal for the stop. The fact that it does not reach peak opening at the 'normal' point of the fricative and maintain it until just before the release of the closure gesture, as we would expect if patterns of coordination were absolute and unchanging, shows that there are no a priori assumptions which we can make regarding the coordination of gestures. We cannot say that glottal gestures coordinate in X-manner with closure or critical gestures without taking account of the overall environment.

Glottal opening gestures are rare cross linguistically if the vocal tract is open, so that e.g. voiceless nasals and sonorants are rare segment types. Instead, if the Supralaryngeal tube is open the glottis tends to have its neutral setting for voicing, but as noted earlier it is not possible to see these neutral settings as absolute. Recapping, we saw that the precise height of the velum differs for stops and fricatives, fricatives not requiring the same tight closure of the velum that stops do, but we can also see that it is not the presence of closure gestures as such that demand tight velic closure, as e.g. /l/ does not have the same requirements as /d/ despite both having the same alveolar closure, /l/ having generally a lower velum than /s/ which in turn has a lower velum than /d/. Rather it is the overall value of the vocal tract which controls velum height so that the tighter the overall constriction of the vocal tract the higher the velum, and each different setting can be viewed as neutral only in that it is the setting chosen when no other is specified.[19] In the same way, if the

[18]We can assume that both the closure and critical gestures have an accompanying glottal gesture but that this is levelled by some form of the Obligatory Contour Principle. What is important is that both the critical and the closure gesture are specified as being overlapped by the glottal gesture.

[19]This is a rather more dynamic interpretation of neutral settings than that proposed by Browman & Goldstein (1989), though it seems to be one that is necessary to handle the phonetic facts.

Supralaryngeal tube is open and there is no gesture active in the glottal tube, the glottis will be set for voicing, and this must be taken into account when considering the role of the glottis in nasal segments.

These various facts are important when considering the relationship between the three active gestures in the Gaelic prenasalised stops in (18). None of the AP forms in (19) require the GLO gesture to overlap the VEL gesture in any way, and this is true of all the prenasalised stops in Gaelic, but we still need to be able to account for the patterns seen. We know for any segment containing a glottal opening gesture and an oral closure gesture that the two are to be coordinated in some way. The theory of heads outlined here provides some constraints on the outcome of this coordination but it does not stipulate whether the glottal gesture should be phased with the oral gesture's onset, target, or offset, it being concerned only to create a single event. By the addition of a headed velic gesture to an otherwise simple voiceless stop we create a constraint that there be two events and thus two VTH trees, and the glottal gesture is, in the absence of any information to the contrary, initially free to phase with either one or both of these trees.

We know from /s/ + stop clusters that a single glottal opening gesture shows compromise between two possible types of coordination as each of the component segments are specified as being voiceless, but as we know that nasals in Gaelic and cross-linguistically show a strong tendency towards voicing it would be surprising if the nasal portion of prenasalised stops were phased with a glottal opening gesture, whether showing compromise or not.[20] In addition, normal voiceless stops in Gaelic have coordination of a certain kind with glottal gestures, and there is no a priori reason why any of these patterns of coordination should be different in any way due to the presence of an additional velic gesture. We can assume therefore that glottal opening gestures in Gaelic are coordinated with the release of closure gestures, as they tend to be cross-linguistically, and in addition they are phased with that part of

[20]By phasing I mean to imply that there is a specific relationship between two or more gestures. Gestures may overlap without being phased to each other.

the closure which results in a closed vocal tract, i.e. the domination period of the LIPS gesture and not that of the VEL gesture.

(19)

a.

[ᵐp] VEL^{H} : **open**, LIPS^{H} : **clo**, GLO : **open**

[ⁿt] VEL^{H} : **open**, TT^{H} : **clo**, GLO : **open**

[ᵑk] VEL^{H} : **open**, TB^{H} : **clo**, GLO : **open**

b.

[ᵐpʰ] VEL^{H} : **open**, LIPS^{H} : **clo**, GLO^{H} : **open**

[ⁿtʰ] VEL^{H} : **open**, TT^{H} : **clo**, GLO^{H} : **open**

[ᵑkʰ] VEL^{H} : **open**, TB^{H} : **clo**, GLO^{H} : **open**

In other words, given a choice of where to coordinate, a glottal opening gesture will coordinate with a period of vocal tract closure and not with a period during which the vocal tract is open, and this accounts for both the glottis' phasing with simple stops and with prenasalised stops. There is simply no requirement within AP for the glottal gesture to phase with the velic gesture in any way; the structures in (19) merely tell us that, all things being equal, some voicelessness will be present, but it does not specify that voicelessness accompany all of the other gestures. In addition, the glottis will continue to create a separate event in the aspirates, so that prenasalised aspirates have three events without the need for any additional means of coordination. As Kingston (1990) points out, the glottis has both a proximate and a distal role in its coordination with closure gestures, but these distinctive roles only apply when a closure gesture produces a closed vocal tract, i.e. a stop, so we expect coordination of glottal and closure gestures to reflect this. The glottis has no such function with nasal segments or velic gestures, and the headed structures outlined here allow this to be directly reflected. The general principles which govern the coordination of gestures predict that prenasalised voiceless stops, or prenasalised voiceless segments of any kind, will have voicing during the nasal portion, just as simple nasals do, and in general voicelessness during any oral portion. The headed structures proposed here constrain gestures to coordinate in a small number of ways so that they form a specific number of audio-acoustic events, but these constraints do not replace the general principles of coordination.

While (18) presents an accurate abstract picture of the results of prenasalisation, the actual realisation of prenasalised stops varies from dialect to dialect, and from speaker to speaker, in three main respects. Each differs in terms of the length of the nasal and oral events, whether or not the aspirated/unaspirated contrast in the stops is neutralised, and the extent to which the stop is voiced.[21] For Barra and the southern Outer Hebrides in general, Borgstrøm (1940) reports that for both series of stops it is the oral part that dominates, which we can represent as $\{^{N}C\}$, where {N} represents the nasal portion, {C} the oral. In much of Skye, on the other hand, neither part dominates, i.e. we find {NC}, though we also find the nasal dominating, i.e. $\{N^{C}\}$. While each dialect differs, the differences are clearly not phonologically relevant, that is there is no difference in the individual categorical structures. Similarly, neutralisation of the aspiration contrast due to prenasalisation is also common, as it is in many other environments in the same dialects. Shuken (1980) reports that in Harris, one speaker treated both series of stops alike, resulting in a voiced nasal followed by an abrupt shift to a brief period (average 22.5ms) of homorganic voiceless oral closure, and in addition treating all nasal + stop clusters in the same way regardless of the environment. These kind of differences do create different categorical structures, though often these can not be seen as caused solely by prenasalisation as such. More interesting for us is the role of the voicing associated with the nasal gesture.

Although the glottal opening gesture in Gaelic prenasalised stops coordinates with the period of oral closure, this closure is rarely fully voiceless, most dialects showing some degree of voicing. The speaker from Harris mentioned above showed

---

[21]We have to rely mainly on phonological dialect descriptions for our data, as little instrumental work has been done on Gaelic, Shuken (1980) being the only work to my knowledge dealing in any way with nasalisation.

a generally abrupt shift to a period of voiceless closure, but occasionally prenasalisation resulted in an aspirated nasal e.g. [nh], while another speaker showed generally full voicing of the prenasalised unaspirated stops, and usually partially voiced but occasionally fully voiceless reflexes of the aspirated stops. This kind of variation is repeated throughout the Gaelic-speaking areas, so that Borgstrøm (1941) describes the prenasalised aspirated stops of Ross-shire as varying between voiceless, half-voiced and fully voiced, with half-voiced being the rule for the area as a whole, some dialects (e.g. Applecross) favouring voiced, others (e.g. Bracadale) favouring voiceless reflexes, but all showing a good deal of variation.[22]

The kind of variation seen here is obviously not phonological as such, but is better handled in terms of overlap and diminution - they are gradient changes which at extremes can appear to give rise to categorical changes. The behaviour of the glottal opening gesture elsewhere in Gaelic provides the key to an accurate analysis of the variations found. The degree of glottal opening seems to be fairly small, even for aspirated stops. I noted earlier that the glottal opening gesture may be larger for unaspirated stops than for aspirated, indicating that size of glottal opening was not a reliable indicator of aspiration, and often the glottis fails to open fully at all. Between vowels, /h/ is usually realised as [ɦ] i.e. as breathy voice, or as [hɦ], and even word-initially it tends to be at least partially voiced. Preaspiration, which also consists of a glottal opening gesture but one which is preconsonantal, is realised in the same way.

In gestural terms, breathy voice is simply the cooccurrence of two glottal gestures with constriction degrees of **cri** and **open** respectively, in other words the glottal opening gesture for /h/ between vowels is at least partially overlapped by the critical gestures of the flanking vowels, the two blending to form breathy voice.

---

[22]Unfortunately, it is often difficult to be certain as to the nature of the glottis in prenasalised stops in Gaelic due to ambiguity in the phonological descriptions. For example, Holmer (1938), describing the Gaelic of Argyllshire, notes that prenasalisation results in voiced stops distinguished from the two voiceless series, but he also describes them as 'weaker than the French, but about equal to English g, b, d' (p73). The stops of English are of course generally only partially voiced, so it is impossible to be certain from the description given whether the stops in this dialect are fully or partially voiced.

While the movement from a GLO : **open** gesture to a GLO : **cri** gesture is abrupt, indicating fast closure, the speed of the reverse movement, that is glottal opening, is much slower. We can see this clearly in the contrast between pre- and postaspiration, where /hp/ in Harris is typically realised as [ɦhp] due to slow opening of the glottis, or as [ɦp], but /pʰ/ is realised as [pʰ], postaspiration showing little or no voicing (20ms at most (Shuken 1980)) because of the relatively swift movement from critical to open, while preaspiration is often partially or even entirely voiced. This tendency of glottal opening gestures to be overlapped by critical gestures is also seen when they accompany oral gestures, i.e. voiceless stops and fricatives also tend to show some voicing. If voicing is present it seems to persist, and all voiceless segments in Gaelic tend to be voiced to some extent, even word initially. Shuken reports for the speakers from Harris mentioned above an average of 43.75% of the duration of unaspirated stops showing voicing intervocalically, and 31% word-finally.

This general behaviour of the glottis can be extended to explain the various patterns of prenasalised stops seen in the different dialects. Glottal opening gestures in Gaelic show a general tendency to weakening in all environments, in that both their size and duration are liable to be diminished in certain environments, the degree of diminution varying according to context. Whether alone or in combination with other gestures, glottal opening is rarely fully coextensive with its accompanying oral gestures, and all stops show some degree of voicing, though the canonical gestural settings differ from dialect to dialect and from speaker to speaker, so that the tendency to voice is greater for some than for others and the degree of voicing may vary according to the environment, medial stops being particularly vulnerable. Given that the stop portion of prenasalised stops is generally shorter than that for plain stops, suggesting a smaller than usual glottal opening gesture, and the presence of flanking voiced segments (a nasal on the one side, and a sonorant or a vowel on the other), it is little surprise that glottal opening diminishes to such a degree that it may no longer be detectable. At extremes the glottis will not open at all and voicing will be continuous.

Dixit (1989) reports similar data for voiceless unaspirated stops in Hindi, in particular in the voicing of medial /p/. Glottal opening and closure movements for aspirated stops showed truncation of peak opening and of duration in word-medial as opposed to word initial position. Beckman et al. (1992) speculate that the same reduction would be true of the glottal gesture for unaspirated voiceless stops. The reduction in both would be due to overlap from the surrounding voiced segments, causing at least perception of voice if not complete voicing, depending on the degree of overlap, and in fact Dixit's subjects showed no evidence of any glottal opening for the majority of /p/ tokens. For Gaelic, at one end we have dialects such as Barra where the glottal gesture diminishes in general only to a small degree, prenasalised stops showing only a small amount of voicing, and at the other end Applecross where the glottis shows a large degree of diminution in many environments and hence prenasalised stops tend to show a large amount of voicing. Both dialects, however, show variation, Barra occasionally showing fully voiced stops, and Applecross partially voiced stops, so that the partially voiced stops of Barra and the fully voiced stops of Applecross are simply media resulting from particular canonical settings, settings which can be affected by a number of processes to give apparently categorical shifts.

The whole range of different realisations of prenasalised stops found in Gaelic can be easily handled within AP. Significantly, we can not only say that the different forms are categorically the same, sharing the same structure despite the different outputs, but again we can straightforwardly describe the various gradient realisations in terms of differences in gestural coordination and size. Just as we restrict the phasing of glottal opening gestures to periods of oral closure,[23] we must also independently provide more gradient information as to the gesture's size, speed of opening (i.e. its stiffness) and so on, information which is built into any physical

---

[23]This does not imply that non-stop segments cannot be voiceless, either in Gaelic or cross-linguistically. In fact, preaspiration in Gaelic often gives rise to both voiceless sonorants and voiceless nasals. These however are due to another more specific rule of Gaelic which overrides the restrictions normally placed on the phasing of glottal gestures. In other languages, prenasalisation can result in voiceless nasals (Herbert 1986).

system, which reflects the canonical settings of the language, and which is dependent upon environment.

The chief advantages of this approach are that there is no need to arbitrarily choose one extreme of realisations of prenasalised stops in Gaelic over another as reflecting the categorical goal, removing all other realisations to a separate phonetic component, and that to specify the many types of variation found in a sense costs us nothing. This is seen clearly in the interpretation given to the eclipsed stops of Applecross by Borgstrøm (1940) and Ternes (1973) respectively. Ternes describes both the aspirated and unaspirated stops of Applecross as being voiced by nasalising (or perhaps better, eclipsing) words but still maintaining the distinction of aspiration, resulting in a four way series as in (20). Crucial to Ternes strictly phonemic description is his claim that the nasal portion of prenasalising particles in Applecross has been lost, hence the forms in (20) contrast with those in (16) in terms not only of voicing but also in the absence of any nasality. Borgstrøm however, in a largely phonetic description, describes the same stops as nasalised, although the nasal is 'reduced' and often absent, which would result in the forms in Applecross being identical to those in (19). The voicing of stops would then still be attributable to the nasal[24] even if the nasal were often, or indeed usually, deleted in speech[25] as in Applecross. What Ternes is describing is the norm in which the nasal is not present.

[24]Borgstrøm refers to the prenasalised stops of Applecross and Diurinish as 'fairly voiced', perhaps implying that they are not always fully voiced, or that the voicing is weaker than for other voiced segments. If this were so it would again suggest that the voicing of stops might be due to overlap by a preceding nasal.

[25]It would be interesting to discover whether any velum lowering is detectable even in forms where it is apparently deleted. A similar pattern is reported for East Sutherland Gaelic (Dorian 1978) where forms with nasals contrast with forms where the nasal is apparently deleted. This would again possibly involve diminution of the nasal gesture, rather than outright deletion.

(20)

a.

ball	[pauɫ] n. 'ball'	am ball	[bauɫ] 'the ball'
dall	[tauɫ] n. 'blind person'	an dall	[dauɫ] 'the blind person'
gall	[kauɫ] n. 'foreigner'	an gall	[gauɫ] 'the foreigner'

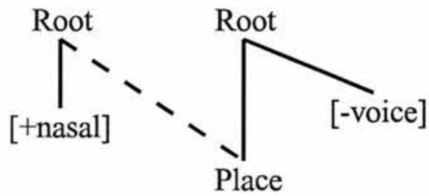
b.

palla	[p^haɫa] n. 'shelf in rock'	am palla	[b^haɫa] 'the shelf'
toll	[t^houɫ] n. 'hole'	an toll	[d^houɫ] 'the hole'
call	[k^hauɫ] n. 'loss'	an call	[g^hauɫ] 'the loss'

In many ways AP treats the facts no differently than does FG, its success lying in the lack of any need for an interface between the categorical and gradient information. Any dialect which showed voicelessness during oral closure would have to be described in FG as in (21), any voicing, even complete voicing, of the oral closure being removed to the phonetics. FG, of course, would not regard this as a problem, as it is not concerned with such fine detail but only in the abstract patterns which arise from the phonetics, especially given that post-nasal voicing of stops in Gaelic can be seen as relatively automatic and mechanical.[26] AP succeeds in that it is capable of describing the same categorical structures as FG, and at the same time it can describe all the various gradient or phonetic realisations in terms of the same basic structures. FG, on the other hand, would require a separate set of secondary mapping rules to map the phonology onto the level of the phonetics. Such variation in voicing as seen in Gaelic is better handled, I believe, within AP where even the finest phonetic details are simply part of the overall phonology.

[26]The alternative would be to represent the voiced and voiceless realisations as being different categorial structures, but this would of course ignore the gradient nature of the realisations.

(21)



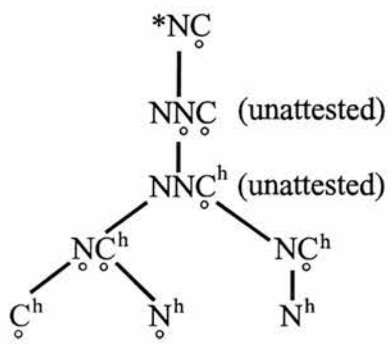
Although the voiced stops in the dialects of Gaelic discussed here might all be attributable to voicing carry-over from a preceding nasal, many languages do show post-nasal voicing as a categorical change (Herbert 1986 claims that it is probably the commonest process undergone by post-nasal consonants), some even showing neutralisation of all voicing contrasts by voicing all prenasalised consonants, e.g. Kamba. There are in addition a number of other processes triggered by prenasalisation which can be analysed fairly simply within the structures proposed here. Herbert lists the commonest developments of nasal + stop in Bantu, as in (22). The addition of a velic lowering gesture often results in voicing of the stop and, as Herbert points out, this voicing is simply a result, diachronically, of the voicing of the nasal gradually spreading to the following consonant. This post-nasal voicing thus becomes phonologised, and we can possibly see this process taking place in some of the Gaelic dialects discussed above. Phonologisation leads to a simplification whereby any segment containing a velic gesture is automatically voiced in its entirety. The development of aspiration, however, is more complex.

(22)

/N + p, t, k/	→	[p, t, k]
		[ph, th, kh]
		[mph, nth, ŋkh]
		[mb, nd, ŋg]
		[m̥, n̥, ŋ̥]

Synchronically, changes such as /N + p/ → [mph] cannot be described as natural processes - we must simply state that addition of the velic gesture implies the presence of a headed glottal gesture - but diachronically a gestural approach might be able to provide an explanation for the creation of such segments. Givón (1974), cited by Herbert, gives the schema in (23) as a description of the development of postaspirated stops under prenasalisation. In the first stage, a phonetic stage which is unattested, an assimilated nasal is presumed to be partially devoiced by a following voiceless, unaspirated stop. In the second stage, which is also unattested, the period of voiceless nasality must be reinterpreted as aspiration. On the basis that voiceless stops tend to be postaspirated, native speakers must apparently first interpret the voiceless nasal as a period of preaspiration, which is marked, and then metathesise the oral and noisy parts so that the stop becomes postaspirated, a less marked segment. The attested forms then develop from this second stage.

(24)



Givón's model implies that changes such as /N + p/ → [mɦ] or [m̥h] must result from a process involving at least four stages, two of which are unattested and are periods of 'purely phonetic' variation (Herbert 1986, p245), which seems unlikely. In addition, the interpretation of a voiceless nasal as aspiration, and the subsequent metathesis, while still perhaps conceivable, also seem unlikely. AP would suggest a simpler, purely articulatory analysis based on differences in gestural overlap and coordination. We have seen in Chapter 1 how AP handles a number of casual speech process which appear categorical but which are in fact gradient, such as the deletion of the coronal stop in 'perfect memory', or the assimilation of /n/ to [m] in 'seven plus'. While there is a great deal of evidence that simple gestural overlap is the underlying cause of these changes (e.g. Barry 1985, Kohler 1976), such changes often secondarily affect a number of other gestures which are coordinated with the gestures involved in the deletion/assimilation.

For example, in the assimilation in 'seven plus', the LIPS gestures for both /v/ and /p/ overlap the TT gesture for /t/, and the subsequent overlap of LIPS and VEL gestures results in 'seve<u>m</u> plus', but the labial stop consists of a glottal opening gesture in addition to the labial closing gesture. This segment is in turn coordinated with the following /l/ and /ʌ/, and the earlier phasing of /p/ with respect to the preceding gestures will generally result in all of these segments also being produced earlier, as long as the relationships between them remain constant. However, if it were a simple matter of moving all the gestures of 'plus' leftwards, so to speak, we might expect the resulting nasal to be at least partially voiceless, as the labial closing gesture and hence the velic gesture would be overlapped by the glottal opening gesture, and similarly in 'perfect memory' the initial labial gesture in 'memory' would overlap the /t/'s accompanying glottal opening gesture.

The lack of any noticeable voiceless nasality might be explainable solely in terms of the relationship between glottal and velic gestures discussed above,[27] but we see the same problems with assimilation involving only oral consonants, as in 'bad cold' realised as 'ba[gk]old', so that a partial geminate is created where the first half is voiced and the second voiceless. This kind of assimilation appears to be present as a categorical process in Havana Liquid Assimilation (Padgett 1991)[28] as in (24), where only the oral gesture spreads, creating either true or partial geminates.[29] While

---

[27]It is important to remember that to describe a stop as voiceless does not imply that the entire period of closure has to be voiceless. All that is necessary is for some of the period of closure to be voiceless, allowing much of the closure to be overlapped by adjacent voicing, though of course different languages will allow different degrees of overlap.

[28]Padgett (1991) asks whether the forms in (18) should be characterised as phonological spreading *or* as 'some sort of gestural overlap' (p231), implying that overlap would involve a phonetic change only. However, the same processes which create casual speech assimilations - e.g. gestural overlap - clearly underlie the changes here. The absence of assimilation in more careful speech points to overlap as the correct analysis.

[29]Padgett points out that although voicing does not normally spread, the fricatives /f x/, when geminated, are realised as /ff/ and /xx/ rather than */vf/ and */ɣx/. He explains this by claiming that as voiced fricatives are not present underlyingly their creation is blocked by a marking condition *[+voice, -son, +cont]. However, this marking condition somehow does not apply to block the

processes such as the deletion of /t/ in 'perfect' and Havana Liquid Assimilation seem identical, they differ on one level. The assimilation in Havana Spanish is made absolute by coordinating some or all of the gestures of a following obstruent with a specific point of the preceding vowel (or liquid) in a regular fashion so that assimilation does or does not take place. Although this involves overlap and has developed from it, it is not a gradient process. Gestural spread can also be at the segmental level and thus involve all the component gestures of a segment rather than just a subset (cf. the fricatives in (24)). In this case the result might be simply deletion, as in 'perfect memory', or the creation of a true geminate (the structure of geminates is discussed in the following chapter), depending on how the overlapping gesture was coordinated.

(24)

ser bobo	[bb]	'to be foolish'	ser pobre	[bp]	'to be poor'
albanil		'mason'	el pobre		'the poor man'
ver droga	[dd]	'to see a drug'	ser tres	[dt]	'to be three'
tal droga		'such a drug'	el tres		'the three'
purga	[gg]	'purge'	parco	[gk]	'parsimonious'
pulga		'flea'	palco		'flea'

ser fino	[ff]	'to be refined'
el fino		'the refined one'
dar jamón	[xx]	'to give ham'
el jamón		'the ham'

In fact, the actual result of overlapping segments containing more than one gesture relies on a large number of factors, such as the consonants' effect on the

---

creation of /ɣ/ and /β/ by Spirantisation, which applies at the same level. The creation of voiceless /ff/ and /xx/ could instead simply be the result of a marking condition specifying that all geminate fricatives are voiceless. Alternatively, and more likely, we can assume that the glottal opening is surrounding vowel formants, the type of gestures involved and so on (Browman & Goldstein 1992a, Byrd 1992).[30] While such detail is beyond the scope of the discussion in hand, it is important for us to recognise the wide variation inherent in the gestural approach. Just as FG can spread individual features or entire segments, we can see that overlap can involve single gestures or clusters of gestures with differing results, with the difference that spreading in FG is absolute while in AP it can be gradient or categorical, and these same processes will naturally operate during nasal assimilation.

Assuming that diachronically nasal assimilation results from gestural overlap, we can expect different languages to show different kinds and amounts of overlap between oral and velic gestures. While in English nasal assimilation to a stop has little effect on the relationship between the stop's oral and glottal gestures, other languages such as Gaelic do show substantial changes. Increasing overlap can lead to partial or complete diminution of glottal opening gestures, so that in e.g. Xhosa the aspirated clicks lose their aspiration altogether when preceded by a nasal. Tarascan (Foster 1969), cited by Herbert, has both voiceless aspirated and unaspirated stops, and nasal assimilation results in voicing of the unaspirated stops and loss of the aspiration of the aspirated stops, e.g. N + /p/ $\rightarrow$ [mb] and N + /p^h/ $\rightarrow$ [mp]. This again shows the effects of diminution of the glottal opening gesture, so that for the unaspirated stops the glottal opening gesture is completely diminished, while for the aspirated stops the relatively larger glottal opening is (perhaps only diachronically speaking[31]) diminished enough to prevent aspiration being created, but not enough to

larger for the fricative than for the stop, and this, combined with its earlier phasing relative to the critical gesture, might result in the fricative being (auditorily at least) voiceless.

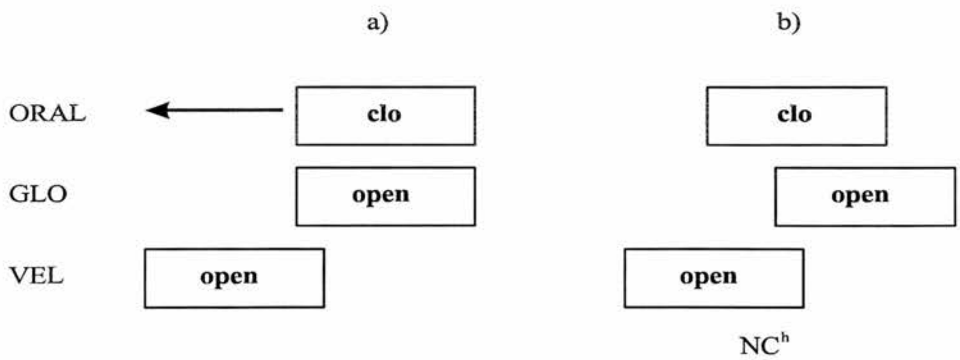
[30]Byrd (1992) describes in some detail the assimilatory effects of increased overlap of C_1 by C_2 in VC_1C_2V sequences, but she only discusses the voiced stops /b, d, g/ so that only oral gestures overlap.

[31]The question of whether the changes in Tarascan are categorial or gradient requires more detailed instrumental investigation. We can compare the final devoicing of obstruents in German and the apparent neutralisation of the voicing contrast. In fact, closer observation shows that the devoiced obstruents remain instrumentally distinct from the underlyingly voiceless (Brockhaus 1991). Similarly, while voicing in prenasalised stops in Tarascan and other languages may be discrete and categorial, it is possibly gradient in the same sense as voicing neutralisation in German and stop voicing in Applecross. This is particularly important in languages which have post-nasal voicing but which do not otherwise have voiced stops.

delete the gesture altogether. Alternatively, glottal opening gestures may be resilient enough to not only remain unaffected by overlapping voicing but to actually devoice the nasal. This is reported synchronically for Amahuaca (Herbert p202), where velic lowering gestures overlap following morpheme-initial obstruents e.g. /wĩ pis/ is realised as [wi $^{\mathring{m}}$pis].

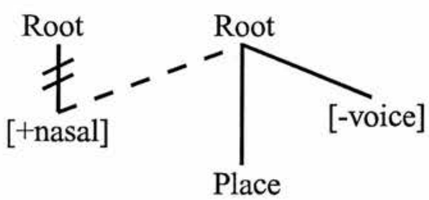
The development of aspiration can be seen as a response to the tendency to voicing and as a means of maintaining voicelessness to some degree. In gestural terms, assimilation of a nasal to a following voiceless stop may be achieved as in (25) where the oral gesture of a voiceless stop moves leftwards to overlap the preceding velic gesture. However, the glottal gesture does not show the same leftward movement, and thus any overlap of the voicing associated with the velic gesture will not result in the diminution nor loss of the glottal gesture. This differs from the type of overlap seen in Havana only in that the coordination of oral and glottal gestures does not remain constant. A side effect of this, as it were, is the creation of postaspiration, which may later be phonologised. This avoids the problems of Givón's model by by-passing the two unattested stages and thus avoiding the need for any metathesis. This is not to say that audio-acoustic factors play no part, but rather that the development of aspiration may have an equally important articulatory origin.

(25)



Depending on the degree of movement of the oral gesture, the assimilation in (25) might result in an aspirated stop, from which ultimately the nasal may be lost, or complete assimilation to an aspirated nasal, as at the foot of Givón's model. The assimilation seen in (16) involves the Root dominating a [+nasal] node, not the [+nasal] node itself, linking directly to a following Place node, so that in some manner the nasal's Root must be aware that it needs to dominate a Place node and so borrows, as it were, the Place node of another segment. Complete assimilation, as in (26), involves a quite different process. Here it is the [nasal] node itself rather than its Root node which assimilates, linking to the following Root node, and thus indirectly to its Place node, and in the process delinking from its own Root node, which is then deleted. The two processes involve assimilation and result in the creation of a nasal segment, but do so by very different means.

(26)



Diachronically, a change from NC to N must involve, in FG, a categorical shift from the assimilation in (16) to that in (26). In AP there is only a change in the velic gesture from a head to a non-head; otherwise the two processes are identical, involving simply the addition of a velic gesture. Complete assimilation can be seen in Lewis Gaelic in (27), where we must assume, for FG, that whereas the adjacent dialect of Harris realises (pre-)nasalisation by assimilation of a Root node to a Place node, Lewis instead assimilates the nasal node directly as in (26). Diachronically we can assume that complete assimilation was achieved by gradually sliding the velic gesture rightwards to overlap the oral gesture, and at the same time gradually diminishing the glottal opening gesture of the stop, though the overlap has proceeded further in Lewis than elsewhere. Synchronically nasalisation in Lewis involves the addition of a non-headed velic gesture, VEL : **open**, as opposed to a headed one, so that to derive /nh/ from /t^h/ involves TT^H : **clo**, GLO^H : **open** becoming TT^H : **clo**, GLO^H : **open**, VEL : **open**.[32] The differences between Harris and Lewis now lie simply in the status of the triggering velic gestures in each, and we do not need to represent them as different types of processes involving different features.

(27)

a.

ball	[pauɫ] n. 'ball'	am ball	[mauɫ] 'the ball'
dall	[tauɫ] n. 'blind person'	an dall	[nauɫ] 'the blind person'
gall	[kauɫ] n. 'foreigner'	an gall	[ŋauɫ] 'the foreigner'

b.

palla	[p^haɫa] n. 'shelf in rock'	am palla	[m^haɫa] 'the shelf'
toll	[t^houɫ] n. 'hole'	an toll	[n^houɫ] 'the hole'
call	[k^hauɫ] n. 'loss'	an call	[$ŋ^h$auɫ] 'the loss'

One point worth noting is that by representing prenasals with two Root nodes, the fact that the target consonant is specified as [-nasal] cannot block assimilation, as it presumably would if prenasals were single segments. However, in the case of complete assimilation as in Lewis this has the unwanted consequence of forcing either stops to be unspecified for [nasal] when assimilation occurs, or for a [-nasal] value to fail to block assimilation and subsequently delete, thus further distancing the processes of Lewis and Harris from each other. This would imply that in nasal spreading processes, [+nasal] could spread to an obstruent, causing deletion of the obstruent's [-nasal] feature, and then continue spreading, resulting in obstruents not only being transparent to nasal spreading but showing full assimilation in the process. I know of no such processes.

---

[32]Compare Goldsmith's (1990) analysis of a very similar process in KiRundi, which involves multiple delinking and relinking, and insertion of Root nodes. None of this is necessary in a gestural approach.

The unaspirated stops also show disjunction between velic and glottal gestures, so that they have either an active velic or glottal opening gesture, but not both. Shuken (1980) shows that even for the aspirated stops, nasalisation results in nasals followed by breathy voice, with no break in voicing from the nasal to the following vowel (resulting in aspiration being strongly nasal breathy voice), again showing the tendency for glottal opening gestures to diminish. Other languages show different patterns. Nasalisation in Sukuma, Pokomo, Ndonga etc. (Herbert 1986) results in voiceless nasals, not showing the disjunction of Lewis Gaelic.[33] Welsh (Ball 1984), with apparently the same process as in Lewis, varies between dialects and individual speakers as to the exact coordination of the glottal and velic gestures, e.g. [nh] ~ [nɦ] ~ [n̥h] ~ [nn̥h]~ [n̥] etc., are all possible realisations of N + /t^h/. While devoicing of the nasal is rare, it does occur, implying either invariant phasing of the glottal gesture with the oral closure gesture as a whole, rather than with that part of it which results in overall closure, or simply that the glottal gesture reaches peak opening early, perhaps before oral closure is achieved.[34] [nh] and [n̥h] differ only in terms of the exact phasing of the glottal gesture: their categorical structures are identical and no language distinguishes between the two, while FG would suggest that they could contrast. There is no additional process of spreading the stop's [laryngeal] features to devoice the nasal for [n̥h], only the normal process of phasing of head and non-head gestures, so that the two show only gradient differences, not the categorical differences claimed by FG.

We can see then how the structures proposed here can provide an adequate representation of nasal segments, and of the processes which create them. The lack of a Place node seems, as Clements (1992) points out, to restrict AP's ability to describe spreading, and similarly the lack of any internal hierarchy between gestures makes the spread of only the oral gestures of a voiceless stop leftwards, without the accompanying glottal gesture, somewhat of an arbitrary process. However, nasal

[33]It is possible that the apparently voiced nasals resulting from nasalisation of non-aspirates are in fact accompanied by a glottal opening gesture, albeit a heavily diminished one.

assimilation as formulated here does not involve spreading, at least synchronically, but simply the addition of a velic gesture, and at the same time the wide variation in realisation is automatically described using the same primitives. The categorical structures are relatively simple, with most of the language particular distinctions arising from gradient differences. Crucially, this gradient information not only specifies the overall velocity, amplitude and so on of individual gestures, but also how they are to be coordinated to match the requirements of their categorical structures, providing at once for the both relative length of the nasal and oral segments of e.g. [^{n}t], but also for the fact that the nasal is voiced rather than voiceless.

In some ways the AP representation is more like single Root theories than the two Root theory seen in (16) in its claim that prenasals are single segments. At the same time, the creation of separate, ordered events in AP with ordered vocal tract hierarchies resembles closely the ordered Roots in two Root theories in FG. The following section examines this apparent paradox and shows how AP is better able to provide an explanation of both the phonological and phonetic behaviour of complex and contour segments.

3.4
Ordering and Unordering
3.4.1
Single Root vs Two Root

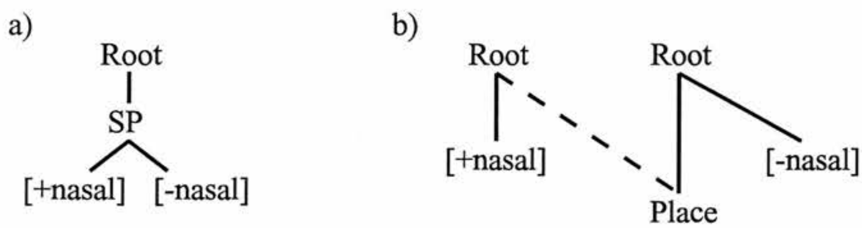
The present view of prenasalised segments in FG stems from the problems raised by Anderson (1976). Anderson noted that a number of languages have segments such as [^{n}t] which phonologically appear to be single segments. However, to treat them as essentially nasal or oral segments fails to capture their true phonological behaviour cross-linguistically and in individual languages, which in

[34]Ball (1984) provides some evidence for early achievement of peak glottal opening for Welsh, noting that intervocalic voiceless stops for some speakers show a small amount of preaspiration. Devoicing of the nasal in [n̥h] would then reflect this coordination of glottal and oral gestures.

many cases is as a sequence of nasal and oral events. As Steriade (1991) points out, prenasalised segments are at once single segments and contain distinct ordered phases of nasality.

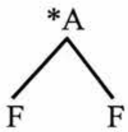
Sagey (1986), following on from Steriade (1982), proposes the (partial) geometry in (28). The nasal/oral nature of prenasals, and their status as single segments, is straightforwardly represented as branching nasal nodes, with two main consequences. Recall that as plus and minus values for a single feature can not be simultaneously realised they must be phonologically ordered, creating a new class of segment which Sagey terms a contour segment (as opposed to complex segments such as /kp/ where there is no phonological ordering, only phonetic). Secondly, if [nasal] can branch in this fashion, we must determine whether other features able to branch in the same manner, and if so what they are and what segments are created. In fact, the number of segments which possibly show such ordering are few in number (affricates being an other oft-cited example), and therefore Sagey proposes to restrict the number and type of branching features by stipulating that only terminal (non-class) features may branch.

(28)



In fact, Sagey's proposal still predicts a larger inventory of contour segments than actually occur in natural languages, and a number of proposals have since been made to restrict their number even further e.g. Piggott (1988), Rosenthall (1989), Selkirk (1991). All of these restrict the number of contour segments by claiming that prenasals are not in fact single segments but a sequence of linked Root nodes ((16), repeated here as (28), preserving the phonological ordering but reclassifying prenasals as clusters. As Padgett (1991) points out, none of these proposals are themselves restrictive enough, and he proposes the condition in (29) which rules out branching of any features. Note that this does not rule out affricates as Padgett follows Lombardi (1990) in assuming that affricates are composed of the privative features [stop] and [cont].

(29)



where F is any feature category
(class or terminal)

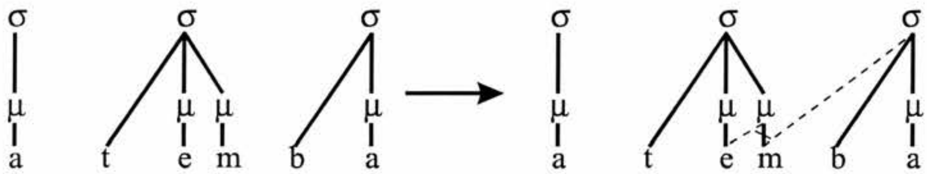
Whatever the precise status of contour segments within FG, two Root theories are claimed to be preferable over single Root for a number of reasons, which I shall discuss below. At the same time there remain two main areas in which single Root theories, whatever their problems, are clearly preferable, areas in which AP and single Root theories appear to agree. These areas rely crucially on prenasals being interpreted as single segments, rather than a sequence of Root nodes which happen to behave as single segments in certain specific environments. If prenasals are in reality simply onset clusters, they are highly marked in that they violate the sonority hierarchy, the nasal being typically more sonorous than the following segment. This finds some support in /s/ + stop onset clusters, which also violate the hierarchy, but remains a problem for two Root theories. More seriously, many languages, e.g. Kikuyu, which otherwise have no onset clusters nevertheless possess prenasalised consonants in onset position. The fact that there are many languages which allow only nasal-obstruent clusters (or other homorganic clusters) is not a solution, as there are many other potential homorganic clusters which never appear as onset clusters.

These problems disappear if we consider prenasals to be single segments. If the prenasals of Kikuyu are single segments then they do not form an exception to the rule of no onset clusters. Similarly, if prenasals are single segments then they do not violate the sonority hierarchy, as it does not apply segment internally. Their component gestures are mutually compatible and coherent in the sense of Steriade

(1991), and in this they resemble other structures which are never considered to be contour segments or clusters, such as aspirated stops or labial-velars. Just as aspirated stops contrast with unaspirated, so simple nasals contrast with prenasals, and as such we should be able to provide a description which reflects this fact.

There are a number of circumstances, however, in which it appears more appropriate to analyse prenasals as clusters, contrary to the predictions of single Root theories. Processes of compensatory lengthening as in Luganda (30) (Maddieson 1993) lend support to the two Root approach. The nasal in the coda resyllabifies rightwards, forming in the process a prenasalised stop, the nasal's mora then freed to link to the preceding vowel. As Padgett (1991) points out, there is no difference between prenasalised stops and nasal-stop clusters: they are both clusters, differing only in syllabification.

(30)



3.4.2

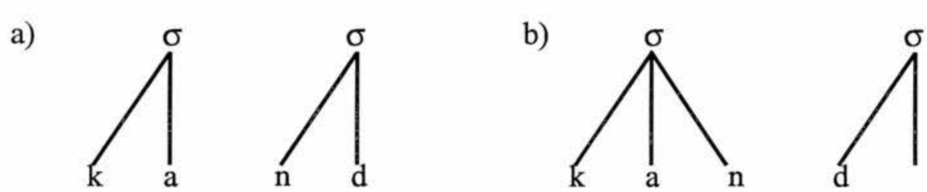
Sinhalese

Within FG, a two Root analysis provides no way to distinguish prenasals from nasal clusters other than by syllabification, whereas single Root analyses of the kind put forward by Sagey predict that the clusters in (28) and the contour segments in (28) are fundamentally different and that languages should be free to contrast them. Padgett discusses the case of Sinhalese, in which there does appear to be a phonological distinction between prenasals and nasal-stop clusters as single Root theories predict. The data in (31), from Feinstein (1979), show a general process of gemination in a class of inanimate nouns to form the definite singular. The forms with prenasals in the plural are clearly distinguished from the nasal-stop clusters in the singular, suggesting a phonological distinction between prenasals and nasal clusters. However, Feinstein suggests that the two forms differ only in terms of their syllabification, so that for singletons the nasal syllabifies rightwards (32) while for geminates the nasal's Root node syllabifies leftwards as in (32). Although there is still no adequate explanation as to why nasals should be unique in showing this ability to syllabify as an onset against the sonority hierarchy, this nevertheless provides an analysis which does not rely on ordered plus and minus values of features. At the same time its prediction that prenasals and nasal clusters differ only in syllabification seems to be borne out.

(31)

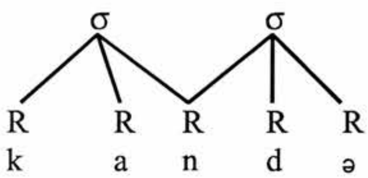
a.			b.		
Sg. Def.	Pl.	Gloss	Sg. Def.	Pl.	Gloss
pottə	potu	core	kandə	ka^{n}du	hill
ginnə	gini	fire	hombə	ho^{m}bu	chin
wattə	watu	estate	hændə	hæ^{n}di	spoon
kæællə	kææli	piece	kondə	ko^{n}du	backbone
reddə	redi	cloth	ændə	æ^{n}di	fence

(32)



Gemination of oral stops in Sinhalese involves the segment's Root node syllabifying both with the following syllable and the preceding, with no delinking. Given this we would initially expect the Root node dominating the nasal to show similar double linking (33), which would be phonetically distinguishable from the prenasal, possibly resulting in a nasal followed by a prenasal cluster.[35] Instead, although the nasal's Root does relink leftwards it also delinks from its onset position as we saw in (32), a move which needs to be stated separately. While maintaining the non-distinction between nasal clusters and prenasals, this is at the expense of needing two rules to describe what appears to be a process requiring only one.

(33)



It is clear from (31) that gemination involves whole segments, so that all the component gestures are syllabified in both the onset and coda.[36] In terms of FG this suggests that a single Root node should dominate all of the features of the segment, and in terms of AP it suggests that there is no difference in the internal, categorical structure of the singletons and geminates of Sinhalese, all differences lying solely in the syllabification. Additional evidence that prenasals are single segments comes from the fact that other than the prenasals, there are no onset clusters in colloquial Sinhalese.[37] We cannot somehow restrict the number or type of gestures which spread, as a number of the segments in (31) contain more than one gesture, and both /n/ and /^{n}d/ contain the same gestures. Browman & Goldstein (1986) suggest that the terminal nodes in the syllabic template in Sinhalese are restricted to oral gestures, and it is these terminal nodes only that are accessed for computing syllable structure. This would then allow prenasals to be syllabified in the onset as they would be invisible to syllable structure rules. However, there is no explanation as to how this

[35]Maddieson (1993) provides an analysis of such a double linking nasal in Sukuma, where both the duration of the nasal and that of the preceding vowel are affected. Double linking in Sinhalese should similarly produce some measurable phonetic reflex, such as a lengthened nasal followed by a shortened oral stop, though of course none such is reported.

[36]I provide some answers as to what onset and coda might refer to in gestural terms in the following chapter. For the moment it is enough to note that such positions exist.

affects the categorical relationship between the oral and velic gestures, precisely where in the syllable structure the velic gestures are located, how this affects other non-oral gestures in simple nasals, laterals and voiceless stops, and what these terminal nodes are or how they are to be integrated into AP. It is clear that it is not just the oral closure gesture which geminates, but the segment as a whole.

We can avoid all of these problems by representing prenasals as single segments. For the class of inanimate nouns in (31) (see Cairns & Feinstein 1982 for other instances of gemination in Sinhalese) we can simply state that for the singular definite the segment as a whole is geminated, that is it is coordinated with both the following and the preceding vowels. As Browman & Goldstein point out, by geminating /t/ to /tt/ we generate a segment that increases in duration as the gestures will be automatically lengthened by being phased with both the preceding and following vowels.[38] The component gestures (TT^H : **clo**, GLO : **open**) do not change in their categorical relationship to each other and still generate a single event. In the same way, the prenasal /^{n}d/ of /ka^{n}du/ and the cluster in /kandə/ have the same internal structure (VEL^H : open, TT^H : clo). Gemination then takes the same form for prenasals as it does for the other segments in (31), with the segment as a whole being coordinated with the following onset and the preceding coda. The sole difference between /^{n}d/ and /n/ then lies in the identity of the velic gesture as a head or a non-head; the two event nature of /^{n}d/ is maintained after gemination, just as the single event nature of the other segments is maintained, but without the need to refer to terminal nodes and producing in the process /nd/.

Single Root theories wrongly predict that there be a distinction between nasal clusters and prenasals, and while two Root theories do not make such a false prediction they do fail to capture the underlying identity of gemination of all segments in Sinhalese. AP, on the other hand, is able to simplify the representation

---

[37]Stop-liquid clusters occur in Sanskrit borrowings in educated speech, but in colloquial speech these clusters are broken up by epenthesis.

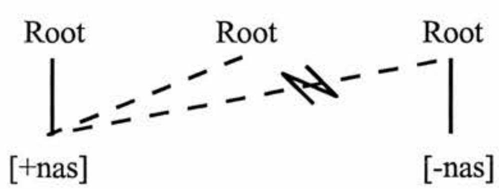
[38]In fact, gemination is a phonologically more complex process than simply altering the coordination between a consonant and its flanking vowels, as I show in chapter 4. However, the theory of gemination set out there would predict the forms found in Sinhalese.

of gemination and at the same time maintain the identity of prenasals as single segments. This meets native speaker intuitions (Cairns & Feinstein 1982), and allows a simpler representation of onsets i.e. there are no onset clusters.

3.4.3
Nasal spreading

There is one remaining phenomenon which causes problems for both the single and two Root theories of FG, and for AP. Padgett (1991) cites Piggott's (1988) claim that nasal spreading rarely results in the creation of prenasals, as further support for two Root approaches. (34) shows the consequences of spreading [+nasal]. Notice that the affected Root node is unspecified for [nasal], allowing the [+nasal] feature to spread. The unaffected Root is however specified [-nasal] and thus blocks spreading for the simple reason that only one value for nasal is allowed (this assumes that no Root node can be interpolated). If we were to assume a single Root approach to prenasals, the presence of a [-nasal] value would not obviously block spreading to it (though it would block spreading past it) as there would be presumably two slots for [nasal] to occupy, and we would require a different stipulation to specifically block the creation of contour segments. Both of these face difficulties when faced with languages such as Terena, where nasal spreading does in fact result in the creation of prenasals.

(34)

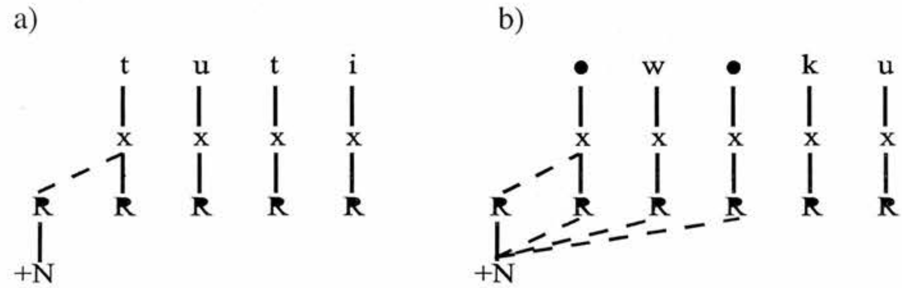


Terena marks the first person by attaching a floating [+nasal] feature to the leftmost non-nasal segment of a word, and then spreads it rightwards, the results of which can be seen in (35). Spreading is blocked by voiceless obstruents, and the obstruent itself is voiced and prenasalised. Although single Root theories can describe the creation of the prenasals very easily - there is an unoccupied [nasal] slot on the obstruent to which the spreading [+nasal] can dock - they also predict that such structures should be far commoner than they in fact are (see Piggott 1988 for other criticisms). Piggott assumes that obstruents and vowels in Terena are unspecified for nasality, and represents the nasal prosody as in (36a) as consisting again of a [+nasal] dominated by a Root node. The Root node then links to the leftmost non-nasal - in this case /t/ - resulting in a prenasalised stop consisting of two Root nodes. A problem arises from this analysis, however. [õwõ$^{\text{ŋ}}$gu] presumably is formed in the same way by linking the Root node of the nasal prosody to the first non-nasal, in this case /o/, and then spreading the prosody itself rightwards unimpeded until the /k/ is reached. This assumes the form in (36b), with the initial vowel now consisting of two Root nodes, but there is nothing to suggest how this might differ from ordinary nasal vowels in its realisation, especially in its ordering implications. Further, this account does not explain the development of word-internal prenasals. Assuming that the nasal prosody linking to the leftmost non-nasal involves the same process in (36a), followed by spreading of the nasal node, this leaves no Root node available to produce the prenasalised /$^{\text{ŋ}}$g/ in [õwõ$^{\text{ŋ}}$gu], unless we assume the insertion of a Root node.

(35)

a.		b.	
emoʔu	'his word'	ẽmõʔũ	'my word'
ayo	'his brother'	ãyõ	'my brother'
owoku	'his house'	õwõ$^{\text{ŋ}}$gu	'my house'
nokone	'his need'	nõ$^{\text{ŋ}}$gone	'my need'
tuti	'his head'	$^{\text{n}}$duti	'my head'
šeʔeša	'his son'	$^{\text{n}}$žeʔeša	'my son'

(36)



AP itself can be subjected to the same criticisms if we insist that nasal spread is a categorical process. Most current phonological theories assume spreading processes are a necessary part of the phonological armoury, either as a fill-in or a feature changing process. As a result, nasal spread will always result in a categorical change in the internal structure of a segment, and will often result in the creation of segments which would not normally be considered part of the underlying inventory. We can place both the nasal vowels and the prenasalised obstruents of Terena in this class. Not all theories agree on this conception, and Bendor & Samuel (1960) propose an analysis of Terena in terms of Firthian prosodic analysis which involves neither feature changing as such, nor the creation of new segments. Instead, the prosody is a property of the word as a whole and does not link to any one particular segment. If, however, we assume that the processes in Terena are categorical in all respects, AP appears to suffer from the same problems as the two Root approach of FG. To create prenasals, any spreading velic gesture must be a head, which gives us two options. If the velic gesture which is added to the initial vowel of e.g. /owoku/ is a headed gesture then we would predict Terena would possess a series of prenasal vowels and glides as well as prenasalised obstruents. On the other hand, the gesture could become headed only when it is added to the obstruent, though we would have to provide an explanation as to how a single gesture can be simultaneously headed and unheaded. Alternatively, the velic gesture could be a non-head, but then we cannot account for the creation of the prenasals.

The problem for AP lies in the representation of nasal spread as a categorical process affecting the internal structure of segments. This view is a result of

Anderson's (1976) analysis noted above and the subsequent reanalysis by Sagey. Whether we represent prenasals as single or two Root structures, we still predict that they should show phonological consequences of ordering, namely edge effects. Ordering results in all prenasalised segments being [+nasal] when viewed from the left hand side and [-nasal], or oral, when viewed from the right. Prenasals will then block nasal spread from any direction (assuming that the [-nasal] value is present when nasal spread takes place), and will themselves cause spreading to the left but not to the right. Sagey (1986) suggests that the pattern of nasal spread in Land Dayak (37) is a result of the presence of underlying prenasals as well as nasals.

(37)

a. mãlu	'strike'	d. sɑmpɛɫ	'extending to'
b. nãbur	'sow'	e. suntɔk	'in need of'
c. ənãk	'child'	f. suŋkoi	'cooked rice'

(37a-c) show [+nasal] spreading rightwards from a simple nasal until blocked by a following supralaryngeal stop, as we would expect, but (37d-e) do not show spreading, despite the presence of a [+nasal] segment. If we assume that the [+nasal] [-nasal] sequence in e.g. /suntɔk/ represents a prenasalised stop, this looks like an instantiation of the edge effects expected from prenasals. However, these underlying prenasals are never present on the surface, becoming instead simple nasals, so that they are only distinguished from normal nasals by the absence of nasal spread. In other words /suntɔk/ surfaces as [sunɔk] and never as *[suntɔk]. This is a serious weakness in the analysis, but there is no mechanism within FG to distinguish between nasals which cause spreading and those that do not other than to represent one as being partially [-nasal]. There seems in Land Dayak to be two different kinds of nasal which differ solely in the behaviour of the velum, and it this difference which we must describe.

Krakow (1989) reports for English word-initial /m/ that the end of velum lowering is roughly synchronous with the end of lip closing, the labial closure then being released and the velum raised. For the forms in (37a-c) we cannot assume this same synchronisation of nasal and oral gestures. The velum here does not rise with the offset of oral closure but instead remains lowered throughout the following vowel. The nasals in (37d-e) on the other hand do seem to show the same pattern as English /m/, with the velum closing at (roughly) the same time as the oral constriction is released. Looked at in this way it would seem that rather than providing a different structure for these forms, a structure which is subsequently altered to create simple nasals, it is the forms in (37a-c) which should receive a different representation. What we should capture is the fact that there is one set of nasals which show the roughly synchronous velum raising and release of oral closure typical of simple nasals, and another set in which the velum remains lowered until it is somehow told to rise by a supralaryngeal consonant. This of course would imply that nasal spreading does not exist as a categorical process.

As noted in chapter 2, Browman & Goldstein (1986) suggest that prenasal stops have ordered velum lowering and raising gestures. Given that simple nasals must similarly lower and raise the velum it is unclear how Browman & Goldstein would distinguish between the two. At the same time this would fail to capture the relationship between simple nasals and prenasals, and would create a complex structure for a small class of segments which arguably includes prenasals as its only member (Padgett 1991). A simpler and more direct explanation would be to suggest that there are two different types of velic gesture possible, rather than the single opening/closing gesture presently permitted. The data in (37) would be easily explainable if we assumed that Land Dayak had both types of velic gesture, one involving lowering and raising i.e. a movement away from and back to a closed position, the other simply lowering the velum. The nasal prosody in Terena would be of the same type. The velic gesture would be added to the initial vowel of /owoku/, and the velum would then simply remain lowered until the /k/ was reached. Although initially it might seem odd to require such additional velic gestures, it simply brings the velum into line with the other articulators.

This analysis assumes, like Bendor & Samuel, that the pattern of nasal spread in Terena is only partially categorical. A VEL : **open** velic gesture will link to the initial vowel of [õwõ^{ŋ}gu], the crucial point being that the gesture involves only lowering of the velum, not raising; the following segments are nasal only because of the long distance effect of lowering the velum. The fact that the velum ultimately raises must be due to the presence of some kind of raising specification in the /k/, which I shall discuss below. There is no prenasalised stop in the sense we have developed here. Instead, I suggest that the nasal is epenthetic in the same way as the nasal stop in French 'thon deux' [tɔ̃ ^{n}dø] (see the discussion below), produced simply by overlap of the velic gesture and the stop's oral gesture. The voicing in Terena prenasals is produced by the same means. Recalling the differences seen in the overlap of velic and oral gestures in Gaelic and the accompanying voicing which could extend through the oral closure to different degrees, we can see that while in Gaelic the various degrees of overlap would generally be regarded as phonetic, the very same processes which in Terena cause voicing of all obstruents are generally seen as phonological. These same processes are also behind the voicing of medial /p/ in Hindi. A contributory factor is doubtless the fact that Terena lacks a contrast for voicing in its obstruent series, so that there is no loss of distinction even if the obstruent is fully voiced (cf. the similar variation in the value of the glottis in those Australian languages which also lack a voicing contrast in their obstruent series (Dixon 1980)). The important point is that these nasalised segments do not themselves contain a velic lowering gesture and no categorically [+nasal] segments are created. In addition we again see the importance of specifying both the actual ordering of gestures and the overlap between them in gradient terms, leaving the categorical structures to express only the fact that some ordering is present.

3.4.4
French

The differences between Terena and Land Dayak highlight two main problems with this analysis which must be addressed. First of all, why and how does nasal spread create prenasalised obstruents in Terena but not in Land Dayak.

Secondly, as noted in the previous chapter, given the underspecification implied in the gestural score, how is it that oral obstruents which apparently contain no velic gestures can cause the velum to close?

Look-ahead models of phonetics (see Cohn 1990 and the references therein) imply that gestures will be activated at the earliest opportunity. In terms of the behaviour of the velum, in a sequence VVN of vowels unspecified for nasality followed by a nasal stop it implies that the velum will begin to lower in an anticipatory fashion at the beginning of the vocalic sequence. The coordinative model used by AP however assumes that gestures have a fairly stable configuration and as such their anticipation field is context-free (Fowler & Saltzman 1993) and a number of studies support this view (e.g. de Jong 1991; Beckman et al. 1992). Therefore gestures can only affect each other where they overlap. We can provide a ready explanation for the ultimate raising of the velum in Terena and Land Dayak. The earlier lowering of the velum in liquids in American speech reported in Vaissière (1988) is similarly due not to anticipation but rather to the lower intrinsic velum height of liquids. Boyce et al. (1990), as we saw in chapter 2, note that an apparent case of anticipation of velum lowering in /lansal/ is in fact due to a lowering of the velum particular to /s/ and there is no anticipation as look-ahead models would predict. The earlier lowering of the velum in liquids in American speech reported in Vaissière (1988) is similarly not due to anticipation but rather to the lower intrinsic velum height of liquids.

By incorporating the suggestions of Mattingly (1990), discussed in the previous chapter, we can provide a ready explanation for the ultimate raising of the velum in Terena and Land Dayak. For example, the labial and glottal gestures for /p/ provide us with information as to the intended articulatory and audio-acoustic demands of the segment, but it also demands that the velum be generally tightly closed, otherwise both the characteristic release burst and the voicelessness would be undermined (Ohala 1993). This tight closure contrasts, as already noted, with the lesser constrictions demanded in turn by fricatives, liquids and glides. Instead we must assume that the constellation of gestures for /p/ demand that the velum be tightly closed. In addition it might also demand a high velocity of velum raising, while /b/ might demand a relatively slower velocity.

It is not only in terms of the neutral settings that segments will differ. Summers (1987) reports that voiced stops differ from voiceless stops in that closing movements of their oral gestures show both smaller velocities and lower jaw displacements. This suggests that the target of the closure gesture of voiceless stops is greater than that for voiced, and we again must attribute this not to the oral gesture as such, but rather to the fact that it a forms larger constellation with a glottal opening gesture. It is this larger combination which then demands tight velic closure and greater displacement in the oral gesture for voiceless stops, such values being specified in the same way for the behaviour of both the velum and the jaw. These values are not universal, and individual languages will choose different values for neutral articulators. Languages may also differ in whether or not they prescribe a setting for an articulator, and in general vowels which do not take part in an oral/nasal contrast will not be specified for nasality. Similarly, glides and liquids may or may not have a neutral setting for the velum, so that Sundanese specifies that the velum must be raised for glides which consequently block nasal spread (though of course the actual height of the velum will be relatively low when compared to the obstruents, and may allow a considerable amount of nasal flow), whereas the closely related Malay does not and glides are transparent to nasal spread (Piggott 1992).

The stops of Terena and Land Dayak then differ in terms perhaps of the speed of closure specified for active velic gestures, and in the speed of the neutral setting of velum closure specified for stops, Land Dayak showing relatively higher values for both, so that in Terena the transition between the velic lowering gesture and the obstruents' velic raising gesture results in a brief nasal segment. In neither case, however, does nasal spread create categorically nasal segments. Allowing constellations of gestures, i.e. segments, to control the neutral settings of the other articulators enables us to directly describe the behaviour of the velum in any given context. We can compare this to the analysis given by Cohn (1990) of the predictions of a look-ahead model. Cohn examines the movement of the velum in

French which, although it does not possess any phonological rules of nasal spreading, does show more coarticulation than, as Cohn describes it, 'needs to be there'.

Look-ahead models make the simple prediction regarding the phonetic transition between unalike feature values that they should show a rapid transition from one value to the other, with little or no interpolation between the two. However, Cohn notes that the transition between [+nasal] and [-nasal] is very much more rapid for [nt] than for [nV], a pattern requiring a phonological explanation. Given that French contains voiced and voiceless stops as well as a series of nasal consonants which contrast for the feature [nasal], the [-cont] obstruents must be underlyingly specified as [-nasal]. Similarly, French contains phonological nasal and oral vowels, so the oral vowels must also be specified as [-nasal]. This is confirmed for the stops at least in (38a,b), where there is a rapid transition into nasality which takes place during the vowel. As the sonorants do not contrast for nasality, Cohn suggests that they should not be specified for [nasal], and if this continued into the phonetics we would expect a sequence /lV/ to show interpolation of nasality through the /l/ so that it would be nasalised to a large extent. The /l/ in (38c), however, behaves similarly to the stops, showing little nasal flow, so Cohn proposes the rule in (39) requiring that all onset segments except nasal stops are specified as [-nasal].[39] As all stops and oral vowels are now specified as [-nasal] underlyingly, and glides and liquids in the onset will be [-nasal] by (39), we expect that all should show the same general pattern of transition between opposing values for [nasal] when in the onset.

(38)

a. thon	/tɔ̃/	'tuna'
b. daim	/dɛ̃/	'deer'
c. long	/lɔ̃/	'long'
d. bonte	/bɔ̃te/	'goodness'

[39](34) is restricted to the onset as glides and liquids in the coda appear to be unspecified for [nasal] in that they show a large degree of overlap.

e. thon d(eux)	/tɔ̃ dø/	'tuna (two)'
f. bon lait	/bɔ̃ lɛ/	'good milk'
g. belle Nel	/bɛl nɛl/	'pretty Nel'
h. Nöel	/nɔel/	'Christmas'

(39)

Syllable Onset Default Rule

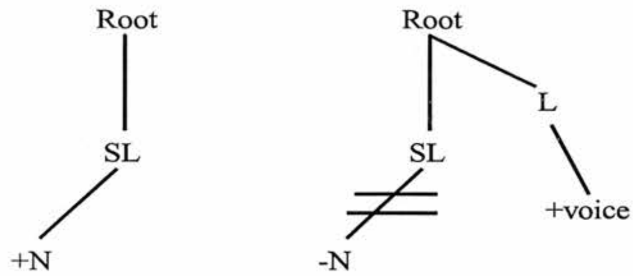
$$\begin{array}{ccc} \text{O} & & \text{O} \\ | & & | \\ \sigma & \longrightarrow & \sigma \\ | & & | \\ \text{ØN} & & \text{-N} \end{array}$$

(38d) shows the expected pattern, the velum showing rapid closure into the oral /t/, and we would expect the same rapid transition in (38e) but we find instead that as much as half of the post-nasal /d/ can be nasalised, resulting in a brief epenthetic nasal stop.[40] The lateral in (38f) shows a similar pattern of gradually decreasing nasal flow, so that the majority of the /l/ is nasalised, although the nasal flow does not continue into the following vowel. Again, (38h) shows the same pattern for a sequence NVV, with the /ɔ/ nasalised, showing decreasing nasal flow which neither carries on into the following vowel nor shows the characteristic plateau of underlyingly nasal vowels. To explain this interpolation of a [+nasal] value through following voiced segments, Cohn proposes the rule in (40) which deletes a specification of [-nasal] from any voiced segment, consonantal or vocalic, which follows a segment which is [+nasal]. Although there is no process of nasal spread, the absence of a specification for nasality predicts that a specification of [+nasal] will be interpolated through a following unspecified segment.

[40] As already noted, this stop is simply a result of gestural overlap. The 'prenasal stop' in Terena [õwõ^{ŋ}gu] is of similar length to that in French, yet only the prenasals in the former are interpreted as categorial objects. I suggest instead that both are due simply to gestural overlap.

(40)

Nasal Deletion Rule



To an extent, Cohn's theory makes the correct predictions, but it also has a number of significant failings. Although the deletion rule in (40) operates on the natural class of all voiced segments - voiced stops, voiced fricatives, sonorants and vowels - there is little phonetically or phonologically to explain why such a rule operates on such a class, other than the imperatives of the look-ahead model. The problem is highlighted when the consonants concerned are flanked on either side by nasal vowels, as in (41). In (41a) the /t/ is flanked by phonologically nasal vowels, and shows almost no overlap of nasality, showing instead rapid transitions as the look-ahead theory and its [+nasal] specification predict. In (41b) the /l/ shows nasality throughout its duration due to interpolation, again as predicted by (40) if it is unspecified for [nasal]. However the /d/ in (41c) shows a very different pattern. It does show some nasal flow but it must also have an oral release. Phonetically of course this is to be expected. As Ohala & Ohala (1991) point out, voiced stops can tolerate a large amount of nasalisation as long as the velum is closed before the release to allow the pressure build up which creates the release burst. Cohn acknowledges this, and suggests that although the /d/ is unspecified for [nasal] there is a purely phonetic constraint to the effect that /d/ must have an oral release.

(41)

a. bon thon /bɔ̃ tɔ̃/ 'good tuna'

b. bon lin /bɔ̃ lɛ̃/ 'good flax'

c. dindon /dɛ̃dɔ̃/ 'turkey'

While the phonetic constraint suggested for /d/ and voiced stops in general is undoubtedly correct, this does not take account of the fact that voiced stops are not special in having this kind of requirement. Voiceless stops, as already noted, require that the velum close early and rapidly to ensure that it retains its voicelessness and release burst, but Cohn interprets the constraints for voiced and voiceless stops differently, assuming that /t/ has a target of continuous oral closure while /d/ has only a kind of 'point' target, requiring only an instantaneous oral release at the offset. Sonorants do not require either tight or rapid velum closure, and can tolerate a low velum so long as this does not overly distort their spectrum (Ohala 1983). To account for only the behaviour of voiced stops in these terms is arbitrary and fails to acknowledge the similar requirements of other segments.

This is, of course, not a fatal problem for Cohn's analysis as it is simply a problem of incompleteness. What it does, though, is serve to highlight the differences between the look-ahead approach and that of AP. AP assumes that the velic lowering gesture, all things being equal, will have a stable and constant duration, and we can assume that its target and offset phase in French have a relatively long duration. In a form such as (38a), we can assume that the vowel has a certain inherent duration specified. In the stream of speech however we know that this duration can be affected by a number of factors, such as stress and its coordination with adjacent segments, which can act to lengthen or shorten it, and these same processes can also affect the overall displacement and velocity of the component gestures (Beckman et al 1992). So, although the vowel may have certain specifications for velum height, velocity of lowering and of offset, other segments may contain conflicting values. In (38d) the following /t/ specifies a rapid and tight closure of the velum, while the vowel requires a lowered velum, and the two velic gestures blend to form a rapid transition to a raised velum during the stop. We can compare this with the overlap of a preceding vowel by voiceless and voiced stops respectively as shown by de Jong (1991). De Jong shows that, as predicted by a gestural model, English vowels are shorter before voiceless stops because the longer gestures of these stops require that they be phased with a slightly earlier portion of the vowel. The offset of the vocalic gesture itself remains fairly constant whether overlapped by a voiceless or a voiced stop, its different observed durations then being due almost solely to the differing amounts of overlap with the following stop. De Jong also notes that the closing movement into the earlier phased voiceless stop's oral gesture has a longer acceleration portion. Earlier phasing of the upward movement of the oral closure arrests the downward movement associated with the vocalic gesture, resulting in a blending of the two which has the effect of shortening the vowel but also delaying the attainment of peak velocity of the stop's closure. This might also be true of French (where otherwise we might expect to see no velum lowering whatsoever during the intended stop) resulting in a very short period of rapidly diminishing nasality during the initial portion of the stop.

What then distinguishes the behaviour of postnasal /t/ and /d/ in French? First let us assume again that each specifies a set of neutral values for the non-active articulators. Secondly, I suggest that we directly encode Ohala's observations on the relationship between velic gestures and following voiced/voiceless stops by specifying a lower neutral value of velum height for /d/, a smaller velocity of closure and possibly a difference in coordination between the oral and velic gestures, so that the velum reaches peak closure at a later point relative to the TT gesture of /d/ than for /t/. Following a nasal as in (38e), the velic gesture of the vowel will phase in the same way with the oral gesture of /d/ as for that of /t/, but the neutral value of the velum for /d/ will allow the vowel's velic gesture to continue for a significant amount of time into the stop's period of closure, resulting in a brief nasal stop. For /l/ the velum will have still lower values for overall height, velocity and so on and will consequently show a greater degree of nasalisation. There is no need for any interpolation of the vowel's nasality through the following consonant, or indeed any changes to the velic gesture. Instead its observed differences in behaviour are artefacts of its relationship to surrounding values for the velum specified by other segments.

This analysis is preferable on a number of accounts. Firstly, we independently know that the different neutral specifications for velum height etc. are needed, so their role in French is not unexpected. Secondly, all of the non-nasal consonants of French are specified as [-nasal] and we no longer require the [-nasal] deletion rule in (40) as the behaviour of the velum falls out naturally from the gestural approach. Thirdly, an independent phonetic constraint requiring oral release for /d/ is no longer required, but is instead implicit in the specification for the velum required by the gestural constellation for /d/.

A number of other aspects of the data are also explainable within this approach. As both /l/ and /d/ are specified as [-nasal] in (38b,c) they should display the same general pattern of transition into the following [+nasal] vowel. However, while the transition from the stop takes place during the vowel, the transition from /l/ takes place during the end of the consonant. Cohn explains the difference in terms of ordered priority constraints (after Holmes et al. 1964), but again it is the constraints placed by the look-ahead approach which demands this further set of phonetic rules. As phonological specifications are categorical and absolute, unlike the gradient nature of phonetic rules, a specification such as [-nasal] should map onto the phonetics in a consistent manner, necessitating the use of phonetic priority constraints when this mapping differs from that expected. Instead, the differences in timing of the transition are just what we expect in a gestural approach.

The problems of the look-ahead model can be seen again in the behaviour of the velum in N V_1V_2 sequences such as Noel /nɔɛl/ 'Christmas'. The /ɔ/ is generally nasalised, but the nasality shows a gradual cline until /ɛ/ shows no nasal flow. This is consistent with the postnasal vowel being unspecified for voicing by (40) and V_2 having a [-nasal] specification.[41] However, Cohn reports that another speaker showed a different pattern, significantly showing continued nasal flow into V_2 and showing slower transitions in general. In terms of the look-ahead theory this continued nasal flow would have to then be due to both V_1 and V_2 being unspecified for [nasal], requiring another rule like (40) to delete V_2's [-nasal] value. It seems simpler to state that this other speaker simply has a larger velic gesture and slower velocity of velic raising, without having to claim that either of the vowels' [-nasal] specifications are deleted by (40). In both cases, the vowels lack any specification for the velum and the velum raising portion of the preceding segment is realised almost without interference for both speakers. There is no 'weakening' of the velic gesture necessary in this account. We simply allow the segments' neutral values for non-active articulators to take part in the phonology

3.4.5
Guaraní

While French does not have phonological spreading of the type found in Terena and Land Dayak, each language makes use of the same basic mechanisms. Nasal harmony does not only spread rightwards, however, and the type of velic gesture found in Terena would be inadequate to describe the pattern of nasality found in Guaraní (van der Hulst & Smith 1982; Kiparsky 1985; Piggott 1988, 1992). Nasal harmony in Guaraní is leftward, so it can not be explained in terms of a simple velic lowering gesture, and it presents a number of problems for standard two Root theories of FG. The forms in (42) show that nasal harmony causes an interchange between simple nasals and prenasalised stops, as can be seen from the behaviour of the negative prefix which appears as both [nõ] and [^{n}do]. These 'nasals' occur as prenasals before an oral sequence, the nature of the preceding sequence being irrelevant, and as simple nasals before a nasal sequence. Previous FG analyses have accounted for the patterns found by spreading both [+nasal] and [-nasal] values, with a number of unfortunate consequences. In particular, spreading of both values for [nasal] fails to account for the fact that [+nasal] sequences always precede [-nasal] ones, and most seriously there must be a distinction between the [-nasal] value which spreads, and that of voiceless obstruents which never spreads and which is somehow invisible to any nasal harmony processes (see Piggot 1992 for a more detailed criticism).

---

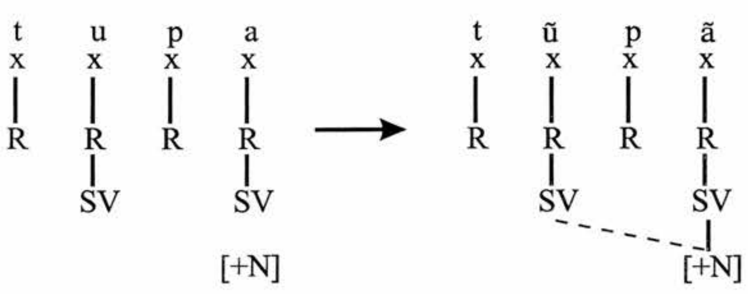
[41]However, the nasal flow is greater through a vowel than through /l/, whereas the look-ahead approach would suggest that both show the same pattern

(42)

a.		b.	
tupa	'bed'	tũpã	'god'
piri	'rush'	pĩrĩ	'to shiver'
haihu	'to love'	mãʔẽ	'to see'
^{m}baʔe	'to be pregnant'	mẽnã	'husband'
^{n}dorohaihui	'I don't love you'	nõrõĩnũpãĩ	'I don't beat you'

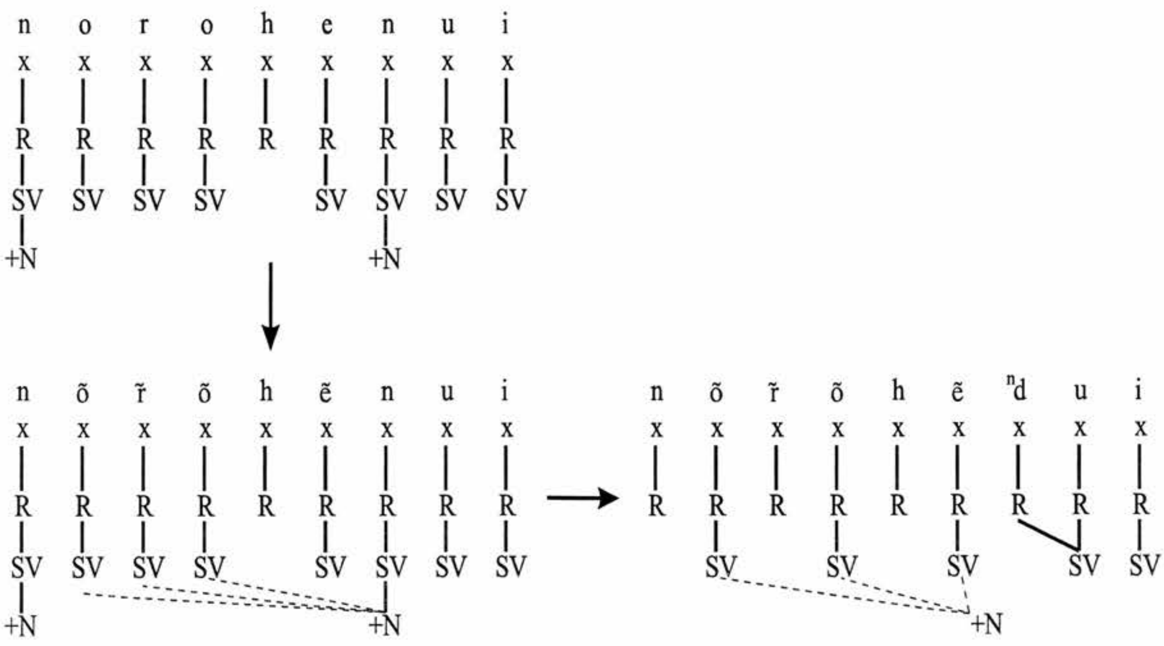
Perhaps the most interesting feature of Guaraní is that all segments, including all obstruents, are transparent to nasal spreading, and Piggott (1992) attempts to account for the differences in transparency and opacity to nasal spreading cross-linguistically by proposing some structural changes to the FG hierarchy. Of concern here is the suggested existence of a Spontaneous Voicing (SV) node and the ability of the nasal node to be dependent either on this or on the more usual Soft Palate node; in Guaraní [nasal] is claimed to be dependent on the SV node. (43) shows his derivation of /tũpã/. Obstruents and laryngeals are unspecified for SV or nasality, while all other segments are specified for SV, with only nasals specified for nasality. Spreading is specified as being leftward and can be caused either by underlying nasals, or by floating [+nasal] autosegments. In (43) the spreading is triggered by a floating [+nasal] feature which attaches to the right edge of the word and subsequently spreads leftwards. The feature links to the final vowel's SV node and thereafter spreads leftwards, targetting other SV nodes, and as obstruents lack such a node they are effectively transparent.

(43)



The derivation above immediately raises questions regarding phonetic implementation. If obstruents receive a specification for nasality before leaving the phonology, so that the intervocalic /p/ in (43) receives a [-nasal] specification, then it is difficult to see how we can insert a [-nasal] feature without causing a delinking of the doubly linked [+nasal]. If obstruents instead remain unspecified for nasality into the phonology, then we would expect the medial /p/ to be realised as either partially or fully nasal. Piggott's account of the nasal-prenasal distinction also suffers from a number of problems of interpretation. In (44) we see Piggott's derivation of [nõr̃õhẽ^{n}dui], where the spreading [+nasal] is lexically associated to the word and thus does not affect the vowels to its right. The harmony rule specifying leftward spread of [+nasal] is not sufficient on its own to explain the appearance of the prenasalised stops. Piggott assumes, along with earlier analyses, that the prenasals derive from underlying nasals which somehow borrow the orality of the following vowel, and this he suggests is due to the Voice Fusion rule in (45).

(44)



(45) Voice Fusion in Guaraní

SV-nodes are fused within a syllable; the features of the right node (i.e. the nucleus or the head of the syllable) dominate

There is nothing about the nature of the SV node which would a priori demand such a process of fusion, either within the syllable or beyond it, so (45) must be specific to Guaraní. Nasal spreading entails the [+nasal] feature of the nasal stop spreading leftwards and linking to preceding SV nodes. Following this (45) will result in delinking of the nasal's SV node and thus of its [+nasal] specification, so that the nasal becomes a non-nasal sonorant stop. This will then be realised as a prenasalised stop by the phonetic implementation rule in (46). The main criticism of this approach is that the nasality of the underlying /n/ and the surface [^{n}d] are unrelated so that the interchange between nasals and prenasals is arbitrary and accidental. This seems unlikely.

(46) The Phonetic Implementation of Spontaneous Voicing

A spontaneously voiced segment contains a nasal phase,
if it is also characterised by complete oral occlusion

I propose instead that it is the prenasals which are underlying in Guaraní and that the nasals are derived, though again this derivation does not involve any categorical change. Gregores & Suárez (1967) note that Guaraní words, including stems plus affixes, are grouped into stress groups which they term macrosegments which consist of a stressed syllable and all preceding unstressed syllables. Nasal spread appears to involve spreading of nasality from wherever it is specified within the word to the beginning of this stress group so that there are no words of the form */tupã/. The various accounts cited above all assume that spreading is a categorical process iteratively spreading [+nasal] leftwards and linking it to other segments; the fact that nasal spread in Guaraní always results in the nasality spreading to the beginning of the word is taken as an 'accidental' consequence of the fact that nothing acts to block spreading. This is a result of the dominant view that spreading must be a categorical process linking features to adjacent segments.

I assume that no matter where the velic gesture is located, its onset is coordinated with the initial vowel of the macrosegment (or more simply word) in a similar way to that of Terena. This directly reflects the observation of Gregores & Suárez that what appears to happen is that the velum begins to lower at the beginning of the macrosegment and progressively lowers until it is at its lowest point at the vowel where the [+nasal] is contrastive. For /tũpã/ this entails coordinating the offset of the velic gesture with the final vowel, as in (43), while its onset is coordinated with the initial vowel. There is no process of spreading velic gestures to other segments in any categorical way.

How then do we account for the non-nasalisation of the obstruents in (43)? Nasal harmony results in the creation of a large velic lowering gesture whose domain is the entire word, and thus initially we would expect this to result in the entire word being characterised by nasality. This assumes, however, that no other segments have any specification for nasality which as we have seen is a false assumption. Guaraní specifies its (voiceless) obstruents as [-nasal], that is the obstruents contain specifications that the velum be raised quickly and held in a tight position throughout most of their duration, just as we saw for French. In /tũpã/ this results in the velum remaining closed during the initial /t/ and then being lowered for the following vowel. The velum will remain lowered until the gestures for the /p/ become active which will again result in the velum raising and then lowering for the following vowel. In /nõr̃õhẽ^{n}dui/ the /r/ does not have any requirements for velum height and thus the velum remains low. However, it is not categorically nasalised in any way. The velic gesture remains active in the same way that the gestures for the vowels are active even when they are overlapped by the consonantal gestures and are inaudible. This characteristic of vowels is made extensive use of in harmony process such as ATR harmony, where consonants are transparent simply because vowels continue to overlap no matter how many consonants intervene. Such harmony does not affect the consonants in any way, and Guaraní nasal spreading behaves in exactly the same way.

Within this analysis the obstruents of Guaraní are not transparent to nasal spread in the manner suggested by Piggott, as the consonants in general have no part to play in the categorical behaviour of the spreading velic gestures. Once the velic gesture coordinates its onset and offset, the consonantal gestures behave in the same general way as those in Terena and French so that the only consonants which are categorically [+nasal] are the underlying nasals. This leaves unanswered the interchange between simple nasals and prenasals, but there are number of facets of the phonetics of the language noted by Gregores & Suárez which suggest that the simple nasals are produced by gestural overlap.

Gregores & Suárez note that the prenasals have two basic forms, [^{N}C] and [N^{C}]. Simply put, prenasals are realised with a dominant oral portion when they are in stressed syllables before an oral span, and in all other environments it is the nasal which dominates, so that initially or medially, if the prenasal syllable is unstressed it will have only a very brief oral release. Before a nasal span of course it has no oral release at all. We can compare Gregores & Suárez's representation of [mẽn^{d}are] 'widower', with a prenasal in an unstressed syllable, with Piggott's [mẽ^{n}dare] where Piggott's semi-phonological transcription somewhat disguises the fact that it is the nasal which dominates.

The voiceless stops /p t k/ themselves are similarly affected by stress, described as having fortis realisation when accompanied by loud stress, but as lenis in other environments. Whatever the precise interpretation given to the fortis-lenis distinction, it is clear that oral stops can undergo various weakening processes, appearing with both reduced displacement and duration. Of further interest is the fact that it is not only the onset of velic gestures which show unusual length, their offsets also appear to be relatively lengthy and persistent, so that following unstressed syllables, even if belonging to an ostensibly oral span, tend to show a fair degree of nasalisation which is only terminated when either a pause or a stressed non-nasal syllable is reached, and this nasality can spread through following liquids and glides e.g. [põr̃ãvã].[42] This results in significant overlap when a strongly stressed nasal syllable precedes an oral one so that an epenthetic nasal stop may be created as in [pẽ tẽ **ĩn** śa] ‘the same’ (stressed syllable in bold), as in French and Terena. Although this only occurs in such strongly stressed positions, it shows that despite their relatively very high specifications for velum height and velocity of closure, the voiceless obstruents can be overlapped sufficiently to produce an epenthetic nasal stop of apparently normal segmental length.

In the underlying form of [nõr̃õhẽ^{n}dui] the initial nasal would have the internal structure typical to a prenasal, with headed velic and oral gestures. These would then be coordinated so that there is only a very brief oral offset. The onset of the second nasal is then coordinated with the beginning of the word so that the velum begins to lower as soon as it can. The velic gesture in [tũpã] behaves in the same way.[43] The voiceless obstruents of Guaraní demand a high velum position with rapid closure, so the lowering movement of the velic gesture is attenuated in [tũpã] so that it does not reach an open position until the constriction for the /t/ is released into the vowel. The initial prenasal in [nõr̃õhẽ^{n}dui] on the other hand has a relatively weak and late velic lowering gesture, as we have seen, with only a very brief oral release, and I suggest that this, combined with preceding and following velic lowering gestures, is sufficient to allow the two lowering gestures to overlap, resulting in the prenasal being realised as a nasal.

The anatomical nature of the vocal tract hierarchy of AP would not allow for the variation in the dependency of [nasal] that Piggot proposes, and there is no way to incorporate such an abstract node as Spontaneous Voicing. The analysis above shows that the behaviour of the velum in Guaraní nasal spreading is of the same kind

---

[42]Gregores & Suárez state that the first and second syllables of [põr̃ãvã] have increasingly strong nasalisation, with slight nasalisation of the final vowel. They also note that the nasalisation does not affect the [v] in 'any noticeable manner' (p66). However, the fact that the final vowel is weakly nasal strongly suggests that the [v] does in fact cooccur with a velic lowering gesture, in the same way as that in [põr̃ãṽãr̃ã], but that the velum is simply relatively high with audible nasality somewhat weakened by the oral airflow.

as that seen elsewhere in the language and cross-linguistically. Just as the offset of velic gestures can persist and spread through following liquids and glides which have no neutral value for the velum, and overlap with a following obstruent to form an epenthetic nasal stop, so the onset can overlap the very short oral portion of a prenasal stop so that it becomes fully nasal. This is the same general process which allows the voiced stops of French and the obstruents of Terena to be nasalised, and no more, and possibly less, overlap is needed in Guaraní. There is no creation of new [+nasal] segments, no SV node fusion, no spreading of [-nasal], and instead of the common nasality of nasals and prenasals being accidental as they are in Piggott's analysis, they are in fact the same objects with no categorical changes whatsoever.

3.4.6
Complex Segments

I return now to the representation in AP of so-called complex segments and their relationship to other types of segment within AP. Recall that Sagey (1986) drew a sharp distinction between contour segments such as /mp/ which showed phonological ordering, and complex segments such as /kp/ which apparently do not show phonological ordering; any phonetic asynchrony in such segments is considered irrelevant. Connell (1994) challenges this view from a number of different perspectives, although, while he highlights the apparent inability of FG to adequately represent these segments, he does not propose any representation of his own.

Connell notes that auditorily, labial-velars resemble labials much more than they do velars, especially in their influence on following vowels. It is in their asynchrony however that he notes the greatest difficulties for FG. As already noted by Maddieson & Ladefoged (1989), the synchrony of complex segments is only impressionistic, and it seems to be the case that the component gestures are always

---

[43]In other words, the gestures for the initial consonant and the spreading velic gesture are all coordinated with roughly the same point of the initial vowel. I discuss consonant to vowel coordination in greater detail in the following chapter.

ordered, typically the TB preceding the LIPS. In addition, while standard assumptions of segmental length in FG would predict that complex segments such as /kp/ should have the same or similar duration to simple /k/ or /p/, Connell notes that in three out of five languages surveyed the labial-velars were significantly longer. This increased duration seems to arise from the fact that they simply involve overlapping TB and LIPS gestures, the overall duration consisting of the inherent duration of the gestures minus the period in which they overlap. Of course, they may overlap to a greater or lesser degree, depending on the language, and Connell notes that in the remaining two languages there is no significant difference in duration between the various segments.

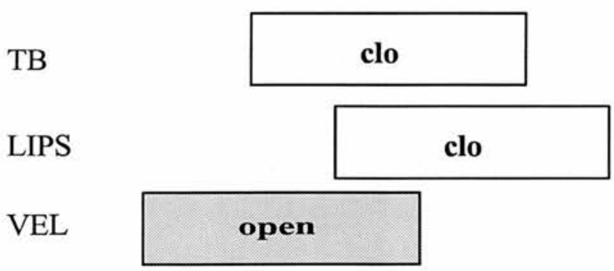
Plausible explanations could possibly be found within FG for these facts, but the result of prenasalising complex segments raises more serious problems which suggest that the notion of phonological unordering and phonetic ordering may be fundamentally wrong. Connell notes that prenasalised complex stops in Ibibio can be regularly realised not as [ŋmkp] but as simple [ŋkp], and that the same pattern is found in other languages such as Efik and Toura (Ohala 1979). Such forms of course are not predicted by FG, in which it is the [place] node which assimilates, and which carries all of its dependants with it. AP, however, can provide an answer.

In the prenasals of Gaelic discussed above we noted how the categorical representation did not provide for the actual coordination of the glottal opening gestures. In this way prenasalisation of voiceless aspirates and unaspirates did not necessitate devoicing of the nasal portion, though this was not ruled out as a possible realisation in other languages. We can compare this directly with the role of the glottal opening gesture in Ibibio /kp/ as described by Connell. Intervocalically /kp/ shows a voicing tail from the preceding vowel so that the initial portion is voiced. In addition, the release is 'prevoiced' so that it has a VOT of -26ms. The internal structure of /kp/ is TB^{H} : **clo**, $LIPS^{H}$: **clo**, GLO : **open**. In some languages the glottal gesture will be phased with the oral gestures so that it is automatically lengthened to produce continuous voicelessness throughout the closure. In Ibibio, however, the glottal gesture will be of the same general length in both simple and complex stops, automatically resulting in the initial and final voicing of /kp/ but leaving its medial portion entirely voiceless.

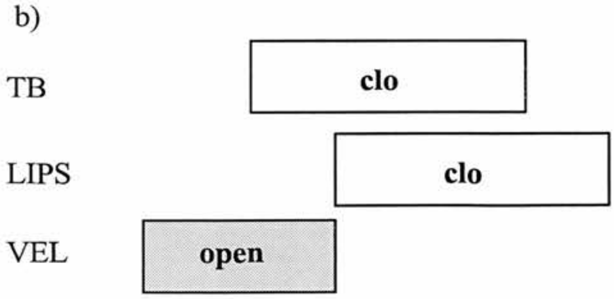
This is analogous to the pattern seen in Gaelic, and we can automatically extend this to cover prenasalised complex stops. Prenasalising /gb/ would involve the addition of a headed velic lowering gesture to give the structure TB^H : **clo**, $LIPS^H$: **clo**, VEL^H : **open**, and we face the same choices here regarding the coordination of the velic gesture as with the coordination of the glottal gesture in /kp/. If the velic gesture were similar to the oral gestures in terms of segmental length, we might expect it to overlap both gestures, but the ultimate output would depend crucially on both the size of all the gestures and the amount of overlap that they show. Just as in Gaelic prenasalised aspirates, there is no a priori requirement for the velic gesture to overlap both oral gestures, merely that there should be some overlap. In (47) we see two possible outcomes of prenasalisation of /gb/, (47a) producing the commoner [ŋmgb], while in (47b) overlap results in [ŋgb]. Both have identical categorical structures, differing only in gestural size and degree of overlap.

(47)

a)



b)



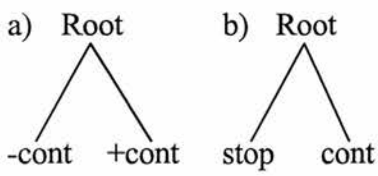
FG could attempt to capture the distinction between the two forms by spreading individual articulator nodes rather than [place], though this would be both arbitrary and would lose the connection between prenasalisation of complex stops and that of simple stops where it is the [place] node which spreads, and would undermine the conception of [place] as an independent node. In addition, by doing this FG would claim that the structures in (47) are categorically distinct and should thus be distinguishable within a single language. Further still, if [dorsal] can spread alone to give [ŋgb], so too should [labial] to give *[mgb], but it does not. AP on the other hand predicts that the observed variation should take place, and that there be no categorical distinctions between the forms in (47). By distinguishing in the way it does between phonological and phonetic ordering, FG is unable to provide an adequate explanation for complex segments. AP, on the other hand, specifies in its categorical structures that some ordering, or better offset, is present, but it does not force a particular ordering on the gestures.

So far we have seen how it is vital to distinguish categorical from gradient information, but also how each is dependent on the other. There is nothing to be gained by placing the two in different domains, connected by some arbitrary mapping parameter. Instead the inherent duration of gestures places equal importance on the actual physical relationships between gestures as well as on the more abstract audio-acoustic goals which these relationships are designed to achieve. There remains one type of segment which has been claimed as a contour segment in FG and elsewhere, but which more recently has been cited as a simple segment, that is affricates. I suggest again that the gestures which constitute an affricate show exactly the same type of categorical structures and gradient interpretation as the other segment types in this chapter.

3.5
Affricates

In terms of production, affricates might best be described as sequences of stops followed by homorganic fricatives. Whether or not these sequences should be interpreted as single segments is a question which has aroused much debate, but most current phonological theories agree that a single segment analysis best accounts for their behaviour cross-linguistically. McCarthy & Prince (1986), for example, show that for correct mapping to the syllable template, Arabic affricates must be analyse as single segments;, representing them as sequences makes incorrect predictions. Both phonetically and in terms of features, however, affricates seem to be identical to sequences of stop plus homorganic fricative, so that e.g. German /ts/ and /t + s/, while phonologically distinct are nevertheless (apparently) phonetically identical. What distinguishes affricates as a segment type is the existence of ordering between the stop and the fricative portions, and Sagey (1986) discussed a number of cases which suggested that this ordering is present at both the phonetic and the phonological levels, as in (48a), where the two values for [cont] must be ordered as they are on the same branch.

(48)



More recently, however, Lombardi (1990) has argued convincingly that this phonological ordering is only apparent. She suggests replacing the binary feature [continuant] with two separate unary features, [cont] replacing [+cont] and [stop] replacing [-cont]. The chief consequence of this move is that the two features [stop] and [cont] are no longer ordered binary features on the same branch and thus there is no phonological ordering. In this way the structure in (48a) becomes that in (48b). If ordering is phonological as Sagey claims, then there should be edge effects, with rules sensitive to the left edge of a segment seeing affricates as stops, and rules sensitive to the right edge seeing them as fricatives. Sagey analysed a number of languages which appear to show such phenomena, but Lombardi provides convincing reinterpretations of these apparent edge effects using unary features, suggesting that ordering of [+/-cont] in each case is purely phonetic (I refer the reader to Lombardi's article for further details). More importantly, and more damaging to the standard analysis, is the existence of data which show apparent anti-edge effects, with affricates behaving as stops in rules sensitive to their right edge, contrary to the predictions of the ordering analysis.

Basque (Hualde 1987; Lombardi 1990; van de Weijer 1993) has a rule deleting stops before other stops, nasals and laterals as in (49a), where stops, nasals and laterals all share the feature [-cont].[44] The simplest description is thus that the first in a sequence [-cont] [-cont] deletes. The affricates in (49b) show the same pattern, deleting their stop portion so that a sequence of underlying affricate-stop surfaces as fricative-stop. In order for the stop portion to delete, however, the affricate's [-cont] node must be adjacent to the [-cont] node of the following segment, and this can only be true if the affricate's opposing values for [cont] are unordered and thus on separate branches. We can continue to refer to plus and minus values for convenience sake, if we wish, but they are now to be interpreted as representing unary features.[45]

(49)

a.

bat paratu	>	ba-paratu	'put one'
guk piztu	>	gu-piztu	'we light'
bat naka	>	banaka	'one by one'
arront lapurre	>	arron-lapurre	'a total thief'

b.

hitz tegi	>	hiztegi	'dictionary'
hitz keta	>	hizketa	'conversation'
haritz ki	>	harizki	'oakwood'

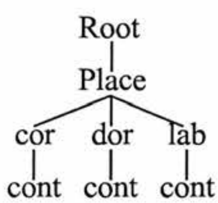
---

[44]The examples are given in Basque orthography. 's/ts' are apico-alveolar /s, ts/; 'z/tz' are predorso-alveolar /c,tc/; 'x/tx' re prepalatal /ʃ, tʃ/.

[45]The precise location of these unary nodes is open to interpretation. Presumably they will either link directly to the Root node, or be grouped under some intermediate Manner node.

AP is, of course, unable to provide an analysis of this type, lacking as it does the necessary abstract Place and Manner nodes. Whether the Manner nodes are binary or unary, or whether they are linked directly to the Root node as in (48a,b) or dependent on the Place nodes as argued by Padgett (1991) (50), the FG analyses hinge on the ability of a single Place node to receive two distinct values for continuancy. Padgett notes a number of problems resulting from the linking of [cont] directly to the Root node, but most serious of these is the inability of such a geometry to describe the various complex segments of Kabardian, e.g. /tx/. Padgett's reanalysis, which is closer at least in spirit to that of AP, itself has problems. Given that complex segments such as /tx/ have two Place nodes dominating two separate values for [cont], and given also the representation of affricates such as /ts/ as having a single Place node alone dominating two separate values for [cont], we might legitimately ask why there are no segments such as */tskx/ with two Place nodes, each dominating two separate values for [cont].

(50)



The problems of FG arise directly form the abstract nature of its features and the accompanying differentiation of complex segments and affricates. In fact, other data from Basque provides evidence that the stop and fricative portions of affricates may only appear to share a single value for Place. Van de Weijer (1993) shows that it is not only the stop portion of the affricate which shows anti-edge effects. The data in (51) show that the second of two adjacent [+cont] nodes deletes, mirroring the deletion of the first of two adjacent [-cont] nodes discussed above. A result of this is that a sequence of fricative-affricate surfaces as fricative-stop (51a). (51b) further shows that when there are two adjacent affricates, both [-cont] deletion and [+cont] deletion are triggered, so that the first affricate loses its [-cont] node while the second loses its [+cont] node, resulting in a surface fricative-stop cluster.

(51)

a.

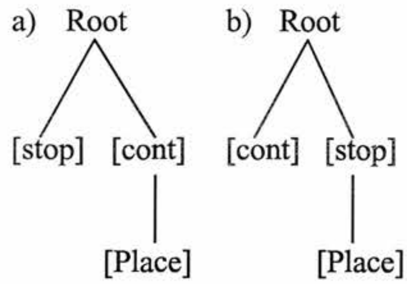
ikas tzen	>	ikasten	'learn' imperative
haz tzen	>	hazten	'grow' imper.

b.

huts tzen	>	husten	'empty' imper.
itx tzen	>	ixten	'close' imper.
utx tzen	>	uxten	'leave' imper.

While this again seems clear evidence that affricates do not have their Manner features ordered, van de Weijer notes that deletion of the affricates' fricative portion always results in the stop /t/ [t̪], no matter what the original Place specification for the affricate, and he proposes the representation in (52a) to explain this. Here there are two unary Manner features, with [cont], the Head, governing the Place node, reflecting van de Weijer's claim that it is the fricative portion of affricates which determines the overall value for Place. Phonetic ordering will universally realise the Head after the non-Head. Van de Weijer further suggests that, at least in Basque, the stop portion will assimilate to the Place specification of the fricative, but that when the fricative i.e. [cont] is deleted, taking with it its specification for Place, the stop will receive a default [coronal] specification. While assimilation of the stop is commonplace - hence the existence of affricates - van de Weijer suggests that it is nevertheless optional, and in (52b), which represents /s/ + stop clusters, the fricative never assimilates, though van de Weijer is unable to say why this should be so.

(52)



Within AP, affricates do not form any special kind of class of segment, but have precisely the same structures as the other segments already described. Given that an affricate such as /ts/ seems clearly to consist of two consecutive ordered events, it seems natural to represent it as containing two separate oral gestures in the same way as e.g. /gb/ or /nd/, so that it has the (partial) internal structure TT^{H} : **clo**, TT^{H} : **cri**. This would result in the tongue tip forming closure and ultimately releasing the closure into a fricative, giving the appearance of a single movement. Unlike FG analyses, there is no shared value for Place. The evidence of Basque, however, seems to suggest that there is in fact no need for such a shared value, and that affricates differ from other segments with more than one gesture solely in that the same articulator set is used for both gestures. Analyses which represent affricates as containing a single specification for Place are typically concerned with Place assimilation e.g. /m/ + /ts/ > /n/ + /ts/. Exactly the same would be achieved by an analysis of /ts/ as containing two separate gestures with assimilation of the nasal to the oral tube.

There are a number of factors which suggest that such an analysis might be correct. The relative frequency of occurrence of affricates in the world's languages as opposed to the comparative rarity of complex segments is seemingly reflected in the fact that they contain a single Place node. In other words, affricates are a less complex type of segment. The same arguments could be applied to AP, though in fact other factors may lie behind the preference for affricates over complex segments. Further, FG must anyway allow different Place nodes to receive different values for

[cont], both in complex segments such as /tx/ and more commonly in consonants with secondary articulation. Further still, phonetically speaking the stop and fricative portions of affricates are often clearly not identical in their specifications for Place, so that in e.g. /pf/ the stop is bilabial but the fricative is labio-dental.

The affricates of Basque strongly suggest that the sharing between the stop and continuant portions of a single value for Place is only apparent. The simplest analysis is one in which the stop portion is identical to the simple stop /t/, while the fricative portions are similarly identical to the simple fricatives. Deletion of the fricative portion would then leave the simple stop behind, without any need for processes of assimilation or default specifications. However, if this were the case, we would surely expect the closure portion of affricates to be dental like that of the simple stop /t/, but this is not so. In fact, given the degree of overlap shown by the gestures of affricates we should expect a good deal of blending of the two values for Place. Both constrictions are formed at roughly the same time - a fact which is more evident for affricates such as /pf/ than for e.g. /ts/ - and the closure gesture will be heavily influenced by the following (and overlapping) critical gesture. For complex segments such as /gb/, the degree of overlap between the two gestures is constrained by the fact that too much overlap of the TB gesture by the LIPS gesture might obscure the velar closure. This is not really a factor for affricates. The target of both the gesture for e.g. /ts/ could in fact be achieved at the same time, with the closure gesture dominating, with no fears that the critical gesture would obscure the closure gesture in any way. While this rules out any possibility of assimilation to a fricative, i.e. assimilation of CD, it suggests that the CL of the fricative portion should dominate, as van de Weijer notes. More specific to Basque is the existence of a number of processes of assimilation of segments to following coronals (Hualde 1991), and a similar type of assimilation may also play a part in the realisation of affricates.

In this way the data in (51) receive a direct and simple explanation, and affricates are no longer a separate segment type but are instead identical to the other segment types discussed in this chapter. Both /pf/ and /ts/ contain two separate gestures, the constriction location of the gestures in the latter being identical and thus appearing to show a single specification. For /tʃ/ there are similarly two separate gestures which have separate values for CL. The fact that both appear to share a single CL is due to the blending we expect in two overlapping gestures, and we can compare this with the superficial identity of the fricatives in 'fi[ʃ] shop' and thi[ʃ] shop' as noted by Local (1991), where the two fricatives are in fact articulatorily distinct. It may be that closer examination may reveal the identity between simple /t/ and the stop portions of Basque fricatives that this analysis suggests.

3.6
Conclusions

The structures proposed here enable us to answer some of the questions posed by various commentators as to the ability of AP to handle certain types of phenomena which do not automatically appear to be amenable to gestural analyses. By using heads in a systematic fashion we are able to clearly distinguish between clusters of gestures which may be viewed as segments and clusters which are better viewed as sequences of segments. In particular, we now no longer need rely on the essentially arbitrary distinction between complete and partial overlap on the one hand and minimal overlap on the other, but can instead automatically generate the necessary overlap between gestures without recourse to these distinctions. Further, as we are no longer reliant on specifying the intergestural relations within segments in terms of coordination of one specific point of a gesture with a specific point of another gesture, there is no longer the possibility as noted by Clements (1992) of generating a far larger number of different segment types than are attested. For example, while the two gestures in a segment such as /gb/ may vary a great deal in terms of the actual overlap between them from language to language or even within a single language, there is only a single structure underlying each of these variants, i.e. TB^{H} : **clo**, $LIPS^{H}$: **clo**. This single structure is constrained in its physical realisation by the demand that there be two distinct and ordered events, but this constraint is satisfied as long as the two gestures are offset regardless as to what form this offsetting actually takes.

Each head, then, controls the overall settings of the vocal tract for a certain period which may or may not coincide with its entire duration, depending on whether another head is also present within the same segment. Non-heads show no such control over the vocal tract. This notion of allowing more than one head, adapted from the conventions of Dependency Phonology, is crucial to the analysis of not only those segments such as /ts/ where there is undoubtedly ordering of some kind between the component gestures, but also segments such as /p^h/ where, at least in FG, ordering is not generally considered to be present. In fact, by allowing segments to have more than one head we give aspirates, prenasals, complex stops and affricates. the same internal structure, the effect of which is to remove the distinction between phonological ordering and phonetic ordering. Ordering between gestures is still present but is not phonologically relevant in the same way.

The actual ordering of the gestures in e.g. /gb/ - both in terms of TB^H : **clo** coming before $LIPS^H$: **clo** and the actual degree of overlap between them - must still be specified in some way, just as it must in any other phonetic model, but where AP has an advantage is in its ability to account for both the discrete, categorical nature of the segment, and its ability to also model the precise physical realisation. It is clearer that the precise ordering cannot automatically follow from the headed structures proposed here when we take into consideration forms such as VEL^H : **open**, TB^H : **clo**. This can, of course, be realised as either /mb/ or /bm/, or variants of these, so that there is nothing inherent in the structure of such segments which compels the nasal portion to be realised before the oral or vice versa. While no language contrasts the two forms phonologically, we obviously cannot provide a single universal ordering, and hence they may be given the same structure.

What remains to be done is to provide similar structures to model the relationships between segments, and it is to this that I now turn in the following chapter.

Chapter 4
Hierarchical Structure and Segmental Length

4
Introduction

The introduction of headed structures allows us to provide the kind of segmentation that has been shown to be necessary by other phonological theories, but these structures provide us with only a local picture of gestural relationships. As Browman & Goldstein (1988) point out, to provide an adequate phonological description we need to capture both the local and the global aspects of gestural coordination. In other words, we must not stop at the segmental level but instead investigate the possibility of the existence of higher structures above this level. For example, if an oral gesture can both stand alone and combine with a velic gesture to produce two distinct segments, it seems reasonable to ask if two such constellations of gestures can further combine to produce yet another level which we could identify with a unit such as an onset or a coda?

One other important question we must ask is how AP can provide an adequate description of segmental length. Clements (1992) points out the difficulty of such a task, and the apparent advantage which more abstract theories have over AP. The principle difficulty lies in the internal duration of gestures. While this is the source of much of the strength of AP as a phonological theory, it also makes it difficult to conceive of any way in which to adequately distinguish between long and short gestures. As Browman & Goldstein (1992b) note, length distinctions should ideally be describable in terms of the same coordinative principles which are seen elsewhere in the theory. There are a number of different approaches to this problem. We can follow fairly closely the views expressed by Clements and posit the existence of some abstract syllabic organisation which would control gestural/segmental coordination and length, or we can view the gestures themselves as always primitive, building up the syllable and length distinctions solely from the relationships between gestures. It is this problem which I shall address here.

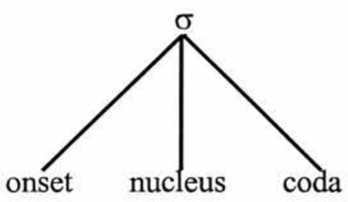
The chapter is organised as follows: In section 4.1 I discuss the general problem of the representation of syllable structure in non-linear phonology; in section 4.2 I discuss the current conceptions of the syllable in AP through the problems posed by Icelandic segmental length and preaspiration; in section 4.3 I outline the solution to the Icelandic data offered by GP and I show how an adaptation of the principles of GP into a gestural framework can provide a fuller analysis of the Icelandic data; in section 4.4 I show how segmental length in Italian requires a different kind of syllable structure than Icelandic; in section 4.5 show how Turkish has two separate kinds of long vowel, identical to those of Icelandic and Italian respectively; section 4.6 provides a summary and pointers to some further areas of research.

4.1
Theories of Syllable Structure

While all current 'phonetic' theories of phonology fail to provide any coherent analysis of syllabic constituency and segmental length contrasts, more traditional abstract analyses appear to be highly successful in many ways, though they are not without their problems. The syllable has been proposed as a phonological unit to simplify the description of a number of areas of phonological structure, providing a simpler analysis of phonotactic constraints and phonological rules. In addition, we find phonological phenomena such as epenthesis and segmental deletion which are often claimed to act to ensure that the phonological string is fully parsed. While we can determine three main conceptions of the syllable within generative phonology, they have much of their basic structure in common.

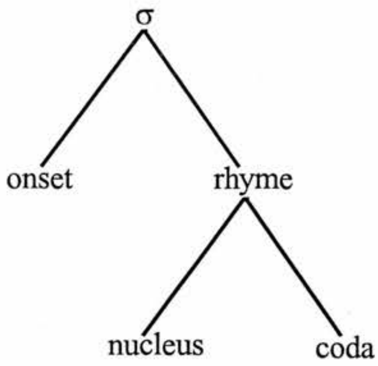
Perhaps the simplest type of structure is that derived from the work of Clements & Keyser (1983) in (1) below. Here we see that three basic units are recognised - onset, nucleus, and coda - though there is no further hierarchical structure, and it is this lack of structure which has attracted the greatest criticism. The flat structure of the tree suggests that each of the constituents should act alone in phonological processes, but on the contrary there seems much evidence to suggest that the constituents often function as if they were part of a hierarchical structure. The evidence for an onset-nucleus grouping rests primarily on the universality of the CV$ syllable (where $ marks a syllable boundary) as opposed to VC$. This universality is itself dependent upon the universal syllabification of a VCV sequence as V$CV, as while languages may lack any words beginning with CV sequences, no language lacks sequences of the type -VCV-. The evidence for a nucleus-coda grouping is perhaps stronger still, particularly in determining overall syllable weight. Neither of these sub-groupings are reflected in (1), and in addition the special status of the nucleus is also lost in this structure, despite its undeniably central role in syllabic organisation.

(1)



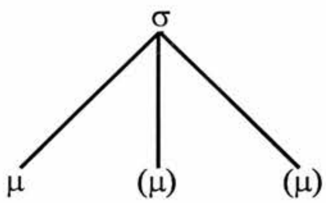
Some of the problems of the flat syllable structure in (1) are addressed by the more elaborate structure in (2). While giving no special status to the nucleus as such, it provides a sub-grouping of the nucleus and coda together under a rhyme node. However, this simply results in a structure of onset versus rhyme and still fails to provide any link between onset and nucleus, though of course this is not a problem if we deny any central status to CV as a basic unit.[1] The one advantage which (2) has is in the creation of the rhyme. In quantity sensitive languages, syllables can be heavy or light, and typically this includes both syllables with long vowels and no coda, and those with short vowel plus coda. Both of these can now be described in terms of branching rhymes (Archangeli 1995b).

(2)



Both (1) and (2), however, largely fail to deal with the problems posed by quantity sensitive languages. In particular, all segments link to the same kind of timing unit, including onsets, yet there is no evidence that onsets should behave in this way. In fact, onsets seem never to contribute to syllable weight and by allowing each segment to associate to a timing unit the view of the rhyme as being the prime bearer of syllable weight becomes arbitrary. The contrast can be seen clearly in Archangeli's (1993) treatment of Yawelmani in terms of the structures in (2), with her (1995a) description of the same data, the latter couched in terms of moraic theory and providing a much simpler and more comprehensive analysis. (3) represents a conception of the syllable in terms of moras (e.g. Hyman 1985; Hayes 1989). Now the identity in weight of long vowels and short vowel plus coda can be seen simply in terms each involving linking to two moras. Again, however, the moraic analysis of syllable structure and segmental length has a number of problems.

(3)[2]



---

[1] It might be argued that we can indeed refer to a basic CV if onset-rhyme is the basic unit and the unmarked rhyme does not branch. This does not, however, provide a *direct* link between onset and nucleus.

[2]The number of moras necessary is open to debate, hence the bracketing in (3). Hayes (1989) suggests that for languages such as Finnish a trimoraic structure is necessary, while for languages with only light syllables only a single mora is needed.

Onsets, being weightless, can link either to the first mora or directly to the syllable node - there is nothing in the theory which predicts this behaviour of the onset, it must simply be stipulated. A consequence of this view of the onset is again the inability to capture the relationship between the onset and the nucleus. Brentari & Bosch (1990) highlight this fact, noting that we still need to refer to the basic CV syllable for processes such as Double Flop in Ancient Greek. In East Ionic dialects /w/ was deleted from onset position e.g. /newos/ 'new' becomes /neos/. In forms such as /odwos/ 'threshold' the same processes applied but resulted in compensatory lengthening of the preceding vowel /o:dos/. Hayes (1989) suggests that following deletion of the /w/ the /d/ delinks from its mora (closed syllables being heavy i.e. bimoraic in Ancient Greek) and 'flops' to the newly free onset position i.e. od$wos to o$dos, and the vowel now spreads to the available mora. However, as 'onset' is not a constituent of the moraic model we cannot state the necessary preference for CV syllables over VC, so the flopping of /d/ to the onset of the following syllable becomes arbitrary.

Brentari & Bosch also note that while moras might be successful in describing quantity sensitive languages, they have no role to play in quantity insensitive languages. Further, weightless or extrametrical consonants are routinely syllabified leftwards and linked to a mora, whether added by rule, or already existing (often removing in the process any link between the mora and the preceding vowel). There seems no reason, however, why such a consonant should not instead simply link directly to the syllable in the same way as onset consonants and thus retain its weightless status. There can be no appeals to the special role of the coda as opposed to the onset as neither are constituents of the theory. In the same way, while moraic theories capture the identity of weight of VV and VC codas in many languages by describing them both as involving two moras, they are unable to account for those languages such as Lardil which distinguish between them, such that only syllables with long vowels count as heavy.

Although each of these theories of syllable structure is successful in many ways, each of them also fails in significant ways, but at the same time they make clear the existence of a number of relationships for which any theory of the syllable must provide an explanation. First of all there must be some central unit which acts as the focus of the syllable, and this we can identify either with the nucleus alone or with the rhyme, the main role of the rhyme being to provide an explanation for syllable weight. Secondly, there is a strong case for CV as the basic syllable type. Finally, all of the theories subsume their constituents under an umbrella node, the syllable node. The extent to which these various aspects should be replicated in AP is an important question.

4.2
Current Theories of the Syllable in AP

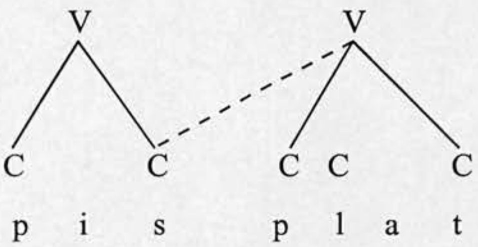
There are then two questions to be answered here. The first is how the various segments are to be coordinated with each other, i.e. vowel to vowel, consonant to vowel ,and consonant to consonant. Secondly, how is it that these relationships give rise to discrete durational differences. Browman & Goldstein (1990) has gone some way to answering the first of these question. They suggest that simple concatenation of gestures can provide at least a first approximation of syllable structure, coordinating gestures (or rather constellations of gestures) in a strictly local manner. However, while this may be adequate for simple words in isolation, in practice it leads to complications.

As we have seen, AP is to some extent an abstraction from the physical data in that it views gestures as critically damped sinusoidal waves which nevertheless have a 360° underlying cycle. From this cycle we can identify a number of points which correspond to perceived discrete portions of any segment. There is general agreement that we can recognise at least three such discrete points in this cycle, corresponding to the onset of the gesture, the achievement of target, and the offset of the gesture i.e. that point at which the tract variables are no longer active in gestural production. We can assume that these same three points are also present in more complex segments such as /gb/ or postaspirates, in other words they are not merely points within the physical production of individual gestures but psychologically as well as physically identifiable points within more complex segments. In this way we can recognise the same points in two different ways for a segment such as /p^h/. Both the LIPS and the glottal gestures will individually contain points which we can mark as onset, target and offset, but at the same time we can see that these same points must be present for the segment as a whole. Onset and offset are readily identifiable with the onset of the LIPS gesture and the offset of the glottal gesture, while the combined target is simply the period from the achievement of the target of the LIPS gesture to the offset of the glottal. While Browman & Goldstein themselves do not discuss this, it is clear that such a view is necessary if we are to move beyond simple gestures to more complex patterns of intergestural coordination.

Assuming this conception of the gestural cycle, we can then coordinate each of these points with a corresponding point in another gesture in a one to one, linear fashion. Browman & Goldstein suggest that the target of all gestures be set at roughly 240° of the underlying cycle, with the offset of consonantal gestures at 290° and of vocalic gestures at 330°. In a CV syllable, the consonant's target will be phased with the onset of the vocalic gesture (i.e. at 0° of the vowel's underlying cycle) so that the vowel is turned on just as the consonant achieves its target. If more than one consonant is present, e.g. CCV, then the vocalic gesture remains coordinated with the leftmost consonant gesture, resulting in the vowel being overlapped to a much greater degree, assuming that each of the consonants belongs to an associated sequence, i.e. they belong to the same syllable as the vowel. There might possibly be some adjustment to the vocalic gesture's stiffness to preserve the overall audible portion of the vowel at a constant duration, but such adjustment is to a large extent arbitrary and not an automatic prediction of the theory. The consonants in a CCV sequence will again be coordinated in a simple concatenative fashion, the onset of the second coinciding with the offset of the preceding consonant. For VC, i.e. nucleus-coda, the pattern is largely the reverse of the CV pattern, the offset of the vowel coordinating with the target of the consonant.

For a C(C)VC sequence, then, the pattern is remarkably similar to the flat syllable structure seen in (1) above, and as such it would be susceptible to the same type of criticisms. Despite the increasing overlap of vocalic gestures as more consonantal gestures are added to the left, there is little to suggest that the vowel-initial consonants here are acting as a phonological unit; they are simply strung together. More serious problems arise when more than one syllable is involved. Browman & Goldstein investigate the sequence /pis plats/ 'piece plots' (Am. Eng.), and suggest that some type of resyllabification takes place between the second vowel /a/ and the final consonant /s/ of the first word. They suggest that the leftmost consonant of *any* intervening consonant sequence coordinates not only not only with the preceding vowel but also with the following, resulting in the structure in (4).

(4)



For a sequence /pi plats/ the onset of the /a/ would coordinate with the target of the immediately preceding /p/, but the addition of /s/ to form /pis plats/ would necessitate a resyllabification so that the onset of /a/ would now coordinate with the target of /s/ rather than /p/. There would now be no specific coordination between /a/ and /p/, and the stiffness of /a/ would again need to be manipulated to prevent it from being obscured and possibly even hidden by increasing overlap from the consonantal gestures. Whether or not such a resyllabification would take place with consonants other than /s/ is a moot point, although it seems highly unlikely that the same processes would take place in sequences such as /pik plats/ or /pil plats/. By allowing such a process Browman & Goldstein propose a structure which is at least one stage less hierarchical than that in (1) as there can now be no real conception of an onset as an independent entity. Further, by allowing such resyllabification we lose any insights into segmental strength and are unable to predict what kind of onset sequences we should find. If any consonants can be strung together in this way there should really be no limit either as to the number of consonants in such a sequence or their identity.

The impetus behind this resyllabification is the desire to ensure that the vowels in each syllable overlap at least minimally. The association of the leftmost consonant in a consonant sequence to both flanking vowels results in those vowels minimally overlapping each other, and this vowel to vowel overlap is an important aspect of the theory.[3] Whatever the number or type of consonants which intervene between two vowels (at least within a word) those vowels will be contiguous and overlap at least minimally. There is much evidence to support this claim. Both Ohman (1966) and Moll & Daniloff (1971) show that the carry-over effects of vocal gestures are long distance. In CVC sequences in American, Swedish and Russian, Ohman noted that the vowels showed both carry-over and anticipatory effects across the intervening consonant. Indeed, coarticulatory effects in general can extend to segments as much as a second away (Benguerel & McFadden 1989). This vowel to vowel coordination has been shown to be present also in gestural models. Silverman & Jun (1993) show that for sequences of /VpkV/ in Korean, where V is either /i/ or /u/, the vowels show clear continuity through the intervening consonants. Zsiga & Byrd (1990) show that the same continuity does not only apply within the word but also across word boundaries.

The complete abandonment of constituency in any form seems too high a price to pay for maintaining vowel to vowel overlap, however, and Browman & Goldstein (1988) redresses some of this imbalance.[4] Browman & Goldstein note that the concatenation of consonants on the basis of the local metric of offset to target probably oversimplifies the types of intergestural relationships actually found, especially since consonant to vowel coordination appears not to be as invariant as was earlier thought. They suggest instead that there exists a global metric, the C-centre, which acts to form an aggregate of word-initial consonants. This C-centre

[3]This vowel to vowel overlap is of course only necessary within words.

[4]Despite the dates, Browman & Goldstein (1988) post-dates Browman & Goldstein (1990) and to a large extent supersedes it.

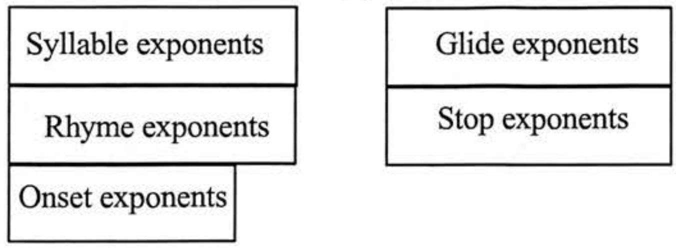
will be coordinated with a fixed point early in the vocalic gesture, so that the entire initial cluster in a word such as 'plot' will be organised so that the central point (of its temporal duration) will extend roughly equally either side of the C-centre. The more consonants present, the shorter the acoustic realisation of the vowel and the longer the overall duration of the syllable. If only a single consonantal gesture is present this will also extend either side of the C-centre.

The advantages of such an organising principle are clear, but it nevertheless entails further reevaluation of the role and identity of the onset, target and offset of gestures. The C-centre provides us with exactly the right type of relationship between the onset and the following nucleus. While consonantal gestures may exist independently of vocalic gestures, the C-centre is specifically defined as being part of the vocalic gesture. In processes such as reduplication, a vocalic gesture may be reduplicated with or without its accompanying onset, but in either case it will still contain a C-centre. This provides us with something approaching an equivalent of the onset, though it leaves open the question as to precisely why this type of coordination is found. At the same time, we can no longer view the onset and so forth of even individual gestures as single points in an underlying cycle. By rejecting the simple concatenative approach and opting for an abstract unit which demands a large degree of overlap between gestures we force ourselves to provide an explanation as to how this overlap occurs.

Coleman (1992) suggests that for stop-glide clusters, rather than concatenate the one with the other we should view them as being overlaid as in (5b), where the phonetic exponents of the /r/ are apparent throughout the /t/. This mirrors the coordination of the onset with its rhyme (5a). While complete overlap is not required for all onset clusters, Coleman explicitly rejects the kind of coordination seen in AP and the symmetrical consonant-vowel transitions arising from it in favour of an asymmetrical model in which the vowel is invariant. The C-centre however moves AP closer to the kind of structures proposed by Coleman, at least in terms of the coordination of gestures within the onset, and it may yet be possible to apply the same principles to the coordination of onset and following vowel.

(5)

a) b)



The segments contained in a C-centre behave to all intents and purposes as if they formed part of a single segment. Rather than simply concatenating the segments of a C-centre we instead coordinate them so that the target portion of each is consecutively realised, in the same way as the segments discussed in chapter 3. The precise amount of overlap which individual segments or gestures may show with others can be regarded as language specific, though constrained by certain physical considerations. In /tr/ all that matters is that both the stop and the sonorant are distinctly realised, while the gestures involved may show more or less overlap. The amount of overlap necessary will vary, depending on the gestures involved, and in /tr/ rather than completely overlapping all of the gestures what we typically find is that the neutral, labial portion of the /r/ extends throughout the cluster while the primary tongue tip gestures show varying degrees of overlap. By coordinating onset segments as if they formed a single segment we can capture both the variation in overlap and the ordering of the individual segments involved.

We can view the C-centre as licensing a certain relationship between consonants, in the sense of Goldsmith (1989, 1990). As well as licensing the appearance of certain consonants and certain sequences of consonants, more importantly for AP it appears to license these consonants to form a macrosegment i.e. a single constituent. Each of the segments in such a cluster, as well as containing at least one headed gesture, will act as a head within this macrosegment, and consequently the same type of coordinative principles apply here as we have seen in the previous chapter. The relationships between the gestures for /pl/ will be fundamentally no different then from the relationships between the gestures in the complex segment /kp/. It is the centre of this macrosegment that will coordinate with the C-centre, that is, the fixed point early in the activation period of the vowel. This is a considerable distance from the earlier concatenative approach, but at the same time it will employ the same basic types of coordination. There is no need to coordinate onset consonants from offset to onset and subsequently compress them in some way to meet the demands of the C-centre, as Browman & Goldstein seem to imply. By specifying that there be discrete events in the onset we imply that the target portions of the respective segments remain distinct, and consequently it is only the target portions which need be actively coordinated. This can be seen most clearly in clusters such as stop + /j/. Languages with such clusters will vary as to the amount of overlap of the two segments, some showing achievement of target for /j/ simultaneous with that of the stop, while others will show only minimal overlap. While this reflects differences in the canonical coordination patterns of each language - presumably through different parts of the target portion being phased - there is no difference in the categorical structures.

Although there appears to be some organising principle influencing the coordination of consonantal gestures to a following associated vocalic gesture, the same does not appear to be true for coda consonants. Browman & Goldstein suggest that while initial consonants do seem to be coordinated according to a global metric, final consonants appear to be coordinated with their associated vowel solely on the local metric of end of activation of the vowel and achievement of target for the consonant. A strict interpretation along these lines however leads to serious problems.

As already noted, Browman & Goldstein investigated the coordination of the various gestures in forms such as /pi(s) p(l)at/, where the bracketed segments may or may not be present. This means that the only syllable-final consonants examined are also word-final, and there is therefore no guarantee that any conclusions reached for these consonants (/s/ and /t/) can be extended to cover word-internal codas. For contrasts such as 'a<u>lt</u>eration' and 'a<u>ll</u> <u>t</u>ogether' there are no generalisations which can be reliably made. Browman & Goldstein claim that the segments in VC$ continue to phase the offset of the vowel with the onset of the consonant, and that these two segments will thus minimally overlap. In a larger sequence of VCC or even VCCC they claim that the subsequent consonants are coordinated with the offset of the preceding segment and that therefore only the first consonant in the sequence will overlap the vowel in any way. This is reflected, they claim, in the putative extrametricality of word-final apical consonants in English words such as 'sixth', and presumably this pattern will be extended to the final consonants in words such as 'cart'. This seems hard to reconcile with forms such as 'carted' where it would seem unlikely that the vowels would not overlap. Such overlap would not be possible however without drastically altering the coordination between the /a/ and the following consonants.

There are more serious consequences of this type of analysis, particularly in Browman & Goldstein's claim that the behaviour of these coda consonants accounts for or correlates with the presumed status of these segments as moraic. This reflects the analysis seen above of syllable final consonants as being heavy. They claim that as these are realised largely 'in their own time' they can somehow be perceived as a timing unit, presumably in addition to the timing unit associated with the preceding vowel. Apart from the question of why the target portion of a word-final consonant, even if produced in its own time frame, should ever be interpreted as equivalent in weight to the preceding vowel, this claim actually goes against the available phonological evidence. As a number of works in the framework of Government Phonology (e.g. Kaye et al 1990) have highlighted, word-final VC syllables seem in fact never to be heavy, contrary to the claims of Browman & Goldstein.[5]

This raises the question as to how precisely segmental length is to be represented. The moraic analysis of coda consonants above does not solve the problem of how to represent long vowels, nor does it provide any link between V: and VC as heavy syllables. Browman & Goldstein note that given their interpretation of codas as only minimally overlapping the preceding vowel, deletion

[5]See Section 2 for further discussion of GP.

of a coda segment would not automatically lead to compensatory lengthening as claimed by Fowler (1983). This would require some other pattern of coordination, one involving much greater overlap between the vowel and its coda consonant, perhaps involving another type of C-centre (Browman & Goldstein 1991). If such differences in coordination exist, they should correspond to differences in phonological structure.

Beckman et al (1992) provide the beginnings of an analysis of segmental length in AP. Examining both phrase-final and stress induced lengthening in English, they show that the lengthening in each is due to two separate processes. Final lengthening appears to be due to an overall slowing down of gestural movements, in other words the speakers examined manipulated the stiffness of the final gestures. Stress induced lengthening on the other hand was created by changing the phasing of the stressed vocalic gesture with the upcoming consonantal gesture (only CVC, and not CV sequences were examined, so it is difficult to know whether the same or similar processes are involved when no consonantal gesture follows). The onset of the following gesture is simply delayed so that less of the vowel is overlapped, automatically resulting in apparent lengthening of the vowel. However, it would be difficult to account for the discrete durational differences found in other languages in either of these ways.[6] Firstly, it would be very difficult to account for the discrete difference between V and V: solely in terms of a change of stiffness. Secondly, to account for length differences in terms of simply delaying the onset of the upcoming gesture and nothing else would seem to be both difficult to constrain and overly powerful.

Whatever method we ultimately choose to represent phonological durational differences in AP, there will necessarily be concomitant changes in gestural stiffness and phasing, but this does not immediately imply that we should directly manipulate these values. The question must arise as to whether the conception of the internal

---

[6] Though as we shall see in the discussion on Italian below, the relationship between a vowel and following segments can account for some types of long vowel, provided we provide a syllabic framework in which this lengthening can take place.

duration of gestures renders AP incapable of providing a phonological account of segmental length. In order to ascertain whether or not this is so, I shall attempt to provide a theory of syllable structure in which all temporal properties are characterised by gestures without recourse to any external timing apparatus other than the internal duration of gestures themselves. In the following section I shall examine the structures of three languages whose syllable structures, while well understood in many ways, nevertheless present a number of problems to current theories of the syllable. An analysis of the structures of these languages within the framework of AP, I suggest, is not only possible but also illuminates areas of these languages which can not be adequately explained in more traditional approaches.

4.3
The Representation of Syllable Structure in a Gestural Framework
4.3.1
Icelandic Syllable Structure

There have been a number of attempts to describe the syllabic structure of Icelandic, with varying degrees of success[7] (e.g. Venneman 1972; Murray & Venneman 1983; Iverson & Kesterson 1989; Thráinsson 1978; Arnasson 1986; Hermans 1984). Indeed, there has so far been little agreement even as to the basic nature of the problem. Primary stress falls in Icelandic on the first syllable of any word and this is reflected in the overall weight of the syllable. (6a) shows that stressed open syllables contain long vowels, assuming that the intervocalic consonants syllabify rightwards into an onset. The same length is also found in (6b). This immediately appears to rule out any possibility of open-syllable lengthening, and to raise the possibility of a trimoraic rhyme. This has generally been rejected, however, in favour of an analysis of the final consonant in each of (6b) as extrasyllabic in some way, preserving the idea of vowel lengthening in open syllables (Venneman 1972).

[7]Gibb (1993) provides strong criticisms of a number of these, to which I refer the reader.

(6)

a.

plata	[pl̥a:ta]	'plate'
tvisvar	[tfi:svar]	'twice'
spara	[spa:ra]	'save'

b.			c.[8]		
rík	[ri:k^h]	'rich'	rikt	[rixt]	'rich' neut.sg.
sæl	[sai:l]	'happy'	sælt	[sæl̥t]	'happy' neut sg.
aum	[au:m]	'miserable'	aumt	[aum̥t]	'miserable' neut.sg.

d.

huss	[huss]	'house' gen. sg.
mann	[man:]	'man'

In (6c-d), however, we see that stressed syllables do not always contain long vowels. In (6c) the adjectives in (6b) are now inflected for the neuter singular by the addition of the suffix /t^h/ which has the effect of shortening the vowel in the stressed syllable. Similarly, (6d) contains forms with final geminates, either derived as in 'húss' or underlying in 'mann', which also contain short vowels in the stressed syllable. Given even the simplest conception of syllabic structure, it is clear that the generalisation that needs to be made is that stressed syllables contain long vowels in open syllables[9] (assuming that e.g. 'aum' contains a final consonant which is extrasyllabic at least until vowel lengthening takes place) but short vowels in closed syllables, and this could be described relatively simply in any of the theories we have seen so far. What is not immediately clear is whether we should be speaking of a process of vowel lengthening in open syllables or vowel shortening in closed syllables.[10] As Arnasson (1986) notes, both would equally well describe the data.

[8]The spirantisation and devoicing will be discussed below.

[9] We shall see below that the generalisation might also be that stressed vowels followed by no more than one consonant are long.

[10]This reflects the generally derivational approaches taken to the problem of Icelandic segmental length.

Additional problems arise when we consider the phenomenon of preaspiration. Icelandic stops (orthographically p,t,k vs b,d,g) are voiceless in all environments and are distinguished solely by the presence or absence of aspiration, i.e. a contrast of /pʰ, tʰ, kʰ/ v /p, t, k/. When these stops are geminated, the non-aspirates undergo no alteration other than one of a change in overall length. Aspirated stops, however, are realised not as simple long stops with a postaspirated release, but as preaspirated simple stops, as in (7). When aspiration is realised to the left of the stop, the stop is no longer postaspirated. The aspiration is realised as a period of breathy voice followed by a much shorter and more variable period of voiceless noise (Kingston 1990), and as such it is directly comparable to the preaspiration in Gaelic seen in chapter 3. The forms in (7) contain underlying geminates, but the same phenomenon is found whether the stops are underlying, or are derived as in (8). The forms in (8b) again involve the neuter singular, and when this is suffixed to a homorganic stop, the resulting cluster is interpreted as a geminate and preaspiration results.[11]

(7)

kappi	[kahpɪ]	'hero'
hattur	[hahtʏr]	'hat'
þakka	[θahka]	'thank'

(8)

a.			b.	
feit	[fei:tʰ]	'fat'	feitt	[feiht]
ljót	[ljou:tʰ]	'ugly'	ljótt	[ljouht]
sæt	[sai:tʰ]	'sweet'	sætt	[saiht]

The important point as far as syllabic structure is concerned is that the aspiration-stop cluster again appears to close the stressed syllable, causing the vowel to be short. The aspiration itself is interpreted as /h/ because of the phonetic identity

[11]Thráinsson (1978) provides a full discussion of all the environments in which preaspiration takes

between the two. In particular, the /h/ in 'hattur' and in 'kappi' both have roughly the same duration, suggesting that preaspiration somehow occupies a timing slot of its own. Preaspiration also takes place in a number of other environments (9) (taken mainly from Jónsson 1994), where it is realised as devoicing of a coda consonant, e.g. South Icelandic [ul̥pa] and not *[ulpa] or *[ulp^{h}a]. In Icelandic the contrast between aspirated and non-aspirated singleton stops is neutralised in a number of environments, generally in favour of non-aspirates in Southern dialects, and in favour of aspirates in Northern dialects. Jónsson (1994) describes Southern dialects as 'resisting' deaspiration only word-initially, so that intervocalically and word-finally after a vowel only non-aspirates are found. In Northern dialects, in contrast, only aspirates are found in these positions. When a sonorant precedes the stop, however, preaspiration again takes place in S Icelandic, realised as devoicing of the sonorant with accompanying loss of postaspiration from the stop. N Icelandic does not share this devoicing.[12]

(9)

	S Icelandic	N Icelandic	
ap^{h}i	a:pi	a:p^{h}i	'monkey'
syrop^h	sy:rop	sy:rop^h	'syrup'
favit^{h}i	fa:viti	fa:vit^{h}i	'idiot'
ulp^{h}a	ul̥pa	ulp^{h}a	'coat'
heimt^{h}a	heim̥ta	heimt^{h}a	'demand'
vant^{h}a	van̥ta	vant^{h}a	'lack'
viŋk^{h}a	viŋ̊ka	viŋk^{h}a	'wave'

Southern Icelandic sonorant devoicing is clearly part of the same general process of aspiration, whereby the aspirate portion of the stop is realised to its left in some environments and to its right in others. The environment in which devoicing

place.

[12]All dialects show devoicing of some kind, so that even in N Icelandic 'maðkur' is realised as [maθkʏr]. There are also a number of 'mixed' dialects in the north which show a mixture of realisations.

occurs is that in which a aspirated onset stop is preceded by a coda, suggesting that it is this particular aspect of syllable structure which is the trigger. The picture is further complicated by the presence of another environment which also triggers preaspiration. In all dialects when an aspirated stop is immediately followed by any one of /l, n, m/ the stop is preaspirated as in (10). The forms in (10a) are lexically underlying, but preaspiration will apply regardless of how the stop and /l n m/ become adjacent. The forms in (10b-c) show the required environment being created both by suffixation of the the genitive plural affix /na/, and by a regular process of syncope. Immediate problems arise as to how these forms should be syllabified. Most (though not all) of these stop + /l n m/ clusters are illicit onset clusters word-initially,[13] and given normal assumptions as to syllable structure they should then be illicit as onset clusters elsewhere in the word. If we allow both the aspiration and the stop to be syllabified into the coda then we have to allow these as the only trimoraic structures in the language. If on the other hand we syllabify the stop into the onset then we create clusters which are otherwise disallowed. Given normal assumptions there is no adequate solution. More significantly, no theory has yet been able to provide an explanation as to why preaspiration should occur in this environment at all.

(10)

a.

epli	[ɛhplɪ]	'apple'
vakna	[vahkna]	'wake up'
ritmi	[rɪhtmɪ]	'rhythm'

b.

pipa	[pʰi:pʰa]	'pipe'	pipna	[pʰihpna]
gata	[ka:tʰa]	'street'	gatna	[kahtna]
kaka	[kʰa:kʰa]	'cake'	kakna	[kʰahkna]

[13]This was not the case in earlier stages of the language.

c.

depill	[tɛ:pɪl]	'dot'	depli	[tɛhplɪ]
jökull	[jœ:kʏl]	'glacier'	jökli	[jœhklɪ]
Bitil	[bɪ:tɪl]	'Beatle' (pop group)	Bitli	[bɪhtlɪ]

Given this alone we might want to state that stressed syllables contain long vowels when followed by a single or no consonant, e.g. 'plata' [pla:tʰa], and short vowels when followed by more than one consonant, e.g. 'kappi' [kahpɪ], 'Bitli' [bɪhtlɪ]. Unfortunately, there are exceptions to this rule. When a stop (or /s/) is followed by any one of /j, v, r/ the vowel remains long, as in (11). Again, while we can account for this by claiming that these clusters are licit onsets word-initially and so presumably licit onsets elsewhere within the word, leaving the stressed syllable open, we might equally well describe this as a process of vowel shortening as one of vowel lengthening, and the question of why preaspiration occurs is left unanswered.

(11)

titra	[tʰi:tr̥a]	'shiver'
vekja	[vɛ:ca]	'awaken'
tvisvar	[tv̥i:svar]	'twice'

Icelandic, then, presents a number of problems. Firstly, we must provide an answer as to how the distinction in vowel length in stressed syllables is created in AP. Secondly, we must show what it is that allows long vowels, however they are represented, to be shortened whenever two or more consonants follow, with the exception of stop + [j, v, r]. Thirdly, and perhaps most importantly, we must show both how preaspiration occurs and why it occurs in such apparently disparate environments. In particular, we must provide a link between the structure of geminates and that of stops followed by [l, n, m]. In the following section I discuss the claims made by Government Phonology, in particular the claims of Gibb (1993), regarding syllable structure, and show how an adaptation of some of the aspects of

GP allows us to provide a gestural description of Icelandic which can be extended to other languages.

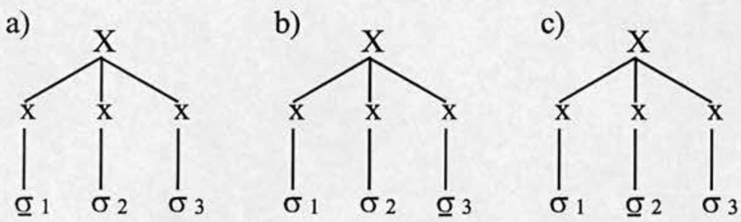
4.3.2
The Analysis of Icelandic in Government Phonology
4.3.2.1
Government Phonology

Government Phonology (GP) attempts to reduce phonology to a single level of representation by removing the need for rules and by eliminating the distinction between underlying and surface structure. While its conception of internal segmental structure differs greatly from that of AP, intersegmental relationships are modelled in a simple binary fashion which can be directly applied to a gestural approach. Both within and between constituents every object must be licensed, or looked at another way, every object must be governed, a view which has extensive implications for syllable structure.

All phonological units, at whatever level, must ultimately be integrated into the prosodic structure if they are to receive phonetic interpretation; failure to be so integrated is in many theories typically seen as a prime motivator of segmental epenthesis, deletion and other phonological processes. In order to integrate these units GP makes use of the priciple of phonological licensing. At the segmental level licensing acts to associate segments to the appropriate skeletal slots, generally constraining which type of segment can associate to which type of slot, what Harris (1994), borrowing from Goldsmith (1990), refers to as autosegmental licensing. At higher levels similar principles apply, ensuring that the skeletal slots and their melodic content are organised into constituents such as the onset, nucleus and rhyme, and on up through the prosodic hierarchy to the foot, word and beyond. In GP this licensing is exhaustive; all phonological units will receive some form of licensing.

The most pervasive form of licensing in GP is government. This is a binary, assymetric relationship which holds between two adjacent skeletal points and is constrained by two main principles, the Adjacency Principle and the Directionality Principle (Kaye et al 1990). The former of these demands that the governor (or head) must be linearly adjacent to its governee (or dependent), while the latter demands that government be strictly directional: within a constituent government goes from left to right, while between constituents it goes from right to left. These conditions constrain all constituents to be maximally binary without having to make binarity a separate condition of syllable structure, as can be seen in (12), where the head of each constituent is underlined. (12a) fails in that node 1 cannot govern node 3 as the two nodes are not strictly adjacent. In (12c) node 2 cannot govern its flanking nodes and observe the principal of directionality. (12b) fails on both counts, as node 3 is not adjacent to node 1, and within the constituent government should be left to right.

(12)



GP recognises three main constituents, namely Onset, Rhyme and Nucleus,where the rhyme is a projection of the nucleus, each of which can form governing domains. These domains are constrained by the Licensing Principle in (13), and we can see the effects of this principle in (14a). Given a nucleus (which projects up to the rhyme) as in (14ai) x_1 is the head of the domain (recalling that within a constituent government goes from left to right) and as such remains unlicensed. In (14aii) the nucleus branches, and x_1 is able to govern x_2, obeying both the Adjacency and the Directionality Principles, and thus license it.[14] In (14aiii) branching occurs at the level of the rhyme but the same principles apply, x_1 governing x_2. That a branching nucleus can not cooccur with a branching rhyme is shown in (14aiv) where the first branching node to govern x_2 does not also govern x_3. The same principles also apply in (14bi, ii), where an onset may or may not branch. In all cases, the fact that x_1 is able to govern x_2 ensures that the latter is fully licensed.

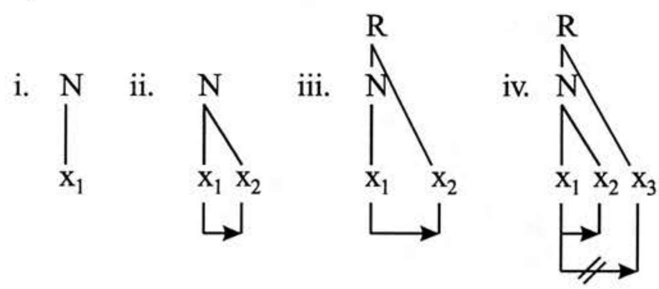
(13)

Licensing Principle

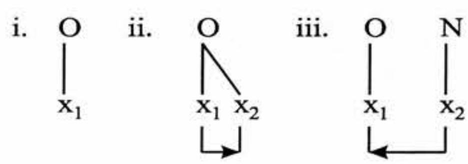
All phonological positions save one must be licensed within a domain. The unlicensed position is the head of this domain.

(14)

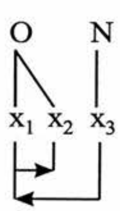
a)



b)



c)



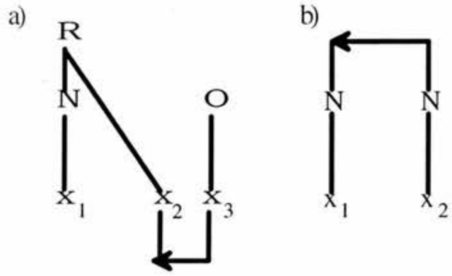
As noted above, government can extend beyond the confines of the constituent to take in interconstituent government. The onset, while an independent constituent, forms part of a larger domain, one involving the onset itself plus a following nucleus. As this is an interconstituent licensing domain it is right headed,

[14]There is no requirement however for constituent heads to govern another position. Constituents may consist of the head without an accompanying dependent.

the nucleus governing the onset and thereby licensing it (14biii). This sets up a dependency whereby all onsets must by definition be followed by a nucleus. We should perhaps recall here that what are being governed are the skeletal slots. When the onset is not branching, the onset head is strictly adjacent to the nucleus head and government is straightforward. However, this is not the case when the onset is branching; instead the two heads are separated by a dependent onset slot (14c). Government is allowed in both instances because x_1 and x_2 form one domain, and x_1 and x_3 another. Charette (1991) proposes a distinction between direct and indirect licensing, the former applying to licensing of a non-branching onset, the latter to the licensing of a branching onset, and the two have different phonological consequences as Charette notes. While there are undoubtedly differences, then, between government within and between constituents, both are generally treated as instances of the same type of government. Leaving this matter aside for the moment, it is clear that while there is no equivalent of the syllable node in GP, the dependency between onsets and following nuclei has the result that all words in GP consist of sequences of onset-rhyme (onset-nucleus) which effectively mimic the syllable. The dependency of the onset on a following rhyme ensures that there are no instances of onsets occurring without an accompanying rhyme.

A branching rhyme (14aiii) is distinct from a branching nucleus (14aii) in that in the fomer the dependent slot is consonantal. However, given the demand that all adjacent skeletal slots show government of some kind or another, we are lead to claim that onsets themselves govern preceding rhymal positions (note that there is no constituent coda in GP; however the rhymal adjunct is usually informally refered to as the coda). This of course means that in order for this rhymal adjunct/coda to be licensed it must be governed by a following onset, therefore if no onset is present, government fails and licensing fails; in other words, the existence of a coda is directly dependent on the presence of a following onset (15a). This would appear to be flatly contradicted by the very large number of languages with words which apparently end in one or more consonants, e.g. English 'foot' or 'child', where there is no overt onset available to license what are usually analysed as codas. Given the licensing claims of GP, a word of the form CVC cannot contain a coda as there is no following onset, therefore the final consonant must be analysed as belonging to an onset. This in turn begs the question of what licenses this onset, as just as all codas are dependent on the existence of following onsets, so all onsets are dependent on following nuclei.

(15)



Rather than create a set of word-final extrametrical consonants which are exempt from the normal principles of licensing, GP claims that such word-final consonants are always followed by a nucleus, though one which is not phonetically realised. Such word-final onset-nucleus clusters are said to contain empty nuclei.[15] These empty nuclei will be phonetically inaudible either if they are properly governed or if they are domain final in a language which parametrically licenses domain final empty nuclei (this is known as the Empty Category Principle). In order for a nucleus to be properly governed it must fulfil the requirements in (16) (Kaye 1990), where proper government is defined in terms of the relationship between two adjacent nuclei as in (15b) above. As in other forms of interconstituent government, in order for one nucleus to govern another they must be adjacent at some level. This adjacency is made possible because the nucleus, as the overall head of the syllable, is able to project itself so that any sequence of nuclei is adjacent at the nucleus projection level whether or not coda or onset consonants intervene.

[15]These nuclei are not in fact completely empty. They instead contain what is known as the cold vowel, which, if realised without the addition of any further vocalic elements, will surface as [ɨ] (see Kaye et al. 1990).

(16)

A nuclear position α properly governs a nuclear position β iff

a. α is adjacent to β on its projection

b. α is not itself licensed

c. No governing domain separates α from β

(17) shows two forms from Moroccan Arabic (Kaye 1990) to illustrate the role of proper government and its effect on empty nuclei. In both [kɨtbu:] and [ktɨb] we have a sequence of three onset-nucleus clusters, each word containing two empty nuclei. In (17a) the final [u:] governs the preceding nucleus allowing it to be phonetically unrealised. However, empty nuclei which are licensed in this way cannot themselves govern a preceding nucleus (by 16b above), and thus the first of the empty nuclei in (17a) cannot be governed and thus must be phonetically realised, resulting in [kɨtbu:]. In (17b) the final empty nuclei is parametrically licensed domain-finally, and again it now cannot govern the preceding empty nucleus which must therefore be phonetically realised. This [ɨ], a full vowel, is now able to govern the preceding empty nuclei which thus remains phonetically unrealised, resulting in [ktɨb].

(17)

a.

O	N	O	N	O	N		[kɨtbu] '(they) write'
x	x	x	x	x	x	x	
k	v^o	t	v^o	b	u		

b.

O	N	O	N	O	N	[ktɨb] 'he writes'
x	x	x	x	x	x	
k	v^o	t	v^o	b	v^o	

As noted above, one of the consequences of this position is that word-final stops never close the preceding syllable so that in 'foot' the /t/ is in fact not part of the preceding syllable but in the onset of a following defective syllable. This accounts for a number of apparently anomalous phenomena associated with such 'codas' (Harris 1994). A further consequence is that rather than inserting epenthetic vowels into syllable structure to make up for defects in syllabification, GP assumes that all vowels which appear on the surface are underlyingly present. While other theories may regard some or all of the instances of [ɨ] in (16) to be underlyingly absent and inserted by rules of epenthesis at varying stages in the syllabification process, GP assumes that they are always present, although they may or may not be phonetically audible.

I show below how an adaptation of much of the structure of GP into AP can allow for a simple theory of syllable structure and segmental length. I suggest that as in GP coda consonants are always dependent upon the presence of a following onset, though this involves a rather different interpretation of what constitutes a coda. Similarly, I suggest an adaptation of GP's analysis of defective syllables and empty vowels, though without GP's insistence on single, unchanging underlying forms. The maximally binary structure of GP, coupled with the licensing of codas by following onsets and the existence of empty vowels, are crucial to Gibb's analysis of Icelandic. However, Gibb's analysis itself has a number of very serious problems, many of which are intrinsic to GP. At the same time however it is the most comprehensive analysis of Icelandic syllable structure and preaspiration, and it raises a number of important points regarding syllabic structure in general which have been largely ignored. I present the details of Gibb's analysis below and then show how the application of a similar analysis within a gestural framework can avoid the problems which GP faces.

4.3.2.2

Icelandic and GP

As we noted above, Icelandic stops are distinguished by the presence versus absence of aspiration, and are voiceless in all environments. Clearly there is something intrinsic to the structure of the aspirated stops as opposed to the non-aspirated which allows aspiration to spread. Features such as [+Spread Glottis] or [+Aspiration] do not obviously account for this. Gibb proposes the representations in (18) for the aspirated stops. The elements of GP which form the internal structure of segments have two distinct properties, depending on whether or not they are combined with other elements and on their status as head or non-head. For example, U^{o} when it stands alone i.e. when combined with no other element, is either the semivowel /w/ or the vowel /u/, depending on its syllabic position. Similarly, I^{o} when syllabified in the nucleus is the vowel /i/, or /j/ when syllabified elsewhere. Every element is free to combine with one or more of the other elements to form more complex head-dependent phonological structures, the results of which combination depending on the internal structure of the segment. Aside from h^{o}, however, only the properties in combination are of concern to us here, as follows: R^{o} contributes coronality, $?^{o}$ can be thought of as 'stoppiness', v^{o} contributes velarity, and H^{-} is stiff vocal cords, indicating a fully voiceless consonant.

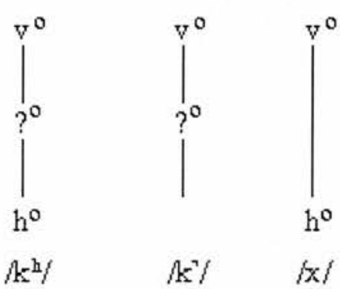
(18)

$$\begin{array}{ccc} U & R & v^{o} \\ | & | & | \\ ?^{o} & ?^{o} & ?^{o} \\ | & | & | \\ H^{-} & H^{-} & H^{-} \\ | & | & | \\ h^{o} & h^{o} & h^{o} \\ /p^{h}/ & /t^{h}/ & /k^{h}/ \end{array}$$

It is the behaviour of h^{o} which provides cause for concern. On its own, h^{o} is realised as /h/, but in combination it makes two slightly different contributions.

When present alongside $?^o$ i.e. in stops, it is realised as the release of that obstruent. Its absence from a stop has been taken to mean absence of release. When combined with elements other than $?^o$ the salient property of h^o is noise or continuancy. This dual role is vital in the GP representation of lenition, as argued by Harris (1990). Compare the contrasts in the segments in (19). Removal of h^o from the velar stop /k/ results in an unreleased [k˺] , as in e.g. 'act', while the removal of $?^o$ results in [x]. Were the segments in (19) to be voiced instead of voiceless, removal of h^o and $?^o$ would result in [g˺] and [ɣ] respectively. As Harris argues, h^o must be present in all released stops in order for lenition to be handled in GP, as it could not be added by any rule because of the general tenet that only elements which are present can spread or be removed - no element can be added to representations if it is not underlyingly present. In other words, if h^o is present in fricatives and released stops can become fricatives by removal of $?^o$, then h^o must be present in released stops.

(19)



Gibb does not provide representations for the voiceless unaspirated stops of Icelandic, but we can make certain assumptions as to their content. As we shall see below, it is the spreading of h^o to adjacent positions which gives rise to aspiration and therefore we might naturally assume it is this element which would distinguish between the two series of stops. Gibb claims that there exists a parameter in Icelandic to the effect that the friction element h^o (at least for stops) must always spread, indirectly implying that unaspirated stops do not contain h^o. Given the arguments of Harris, however, along with the general principle that identical segments should receive identical representations cross-linguistically, all of the stops of Icelandic should contain h^o. If they do not then they will undermine the GP representation of lenition, yet if they do it is difficult to see any reason why h^o should spread from one segment but not the other when it is present in both. Alternatively the stops might be distinguished by H^-, the element which signifies stiff vocal cords, i.e. distinctive voicelessness.[16] We could distinguish the two series by the presence versus absence of H^-, nonaspirates being regarded as having the neutral setting for the larynx and thus being unspecified for H^-. Again, if it is H^- which distinguishes between the two series, why is it not then this element which is spread? There seems no reason why such a difference in the internal structure of aspirated and unaspirated stops should trigger spreading of h^o in one set and not the other.

In terms of syllable structure, Gibb claims that Icelandic has the structures in (20) as templates for well formed onset-nucleus configurations for stressed syllables. We can see that in each of these the rhyme dominates two skeletal points, whether a long vowel or a nucleus-coda sequence. She proposes that Icelandic, amongst other languages, is subject to the Branching Rhyme Constraint (BRC) which is expressed in (21c), the other statements in (21) also being true of Icelandic. Given the division between skeletal or timing slots and the segments which they dominate, the constraints in (21), claims Gibb, not only demand the presence of two skeletal points, but also that they be filled by some material, whether in origin internal or external to the rhyme. These constraints, as we shall see, will not always be met by underlying forms but may rely on a number of repair strategies to arrive at the correct output.

(20)

[16]It is unclear, however, what the realisation of H^- is when unaccompanied.

(21)

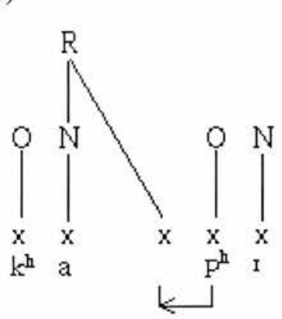
a. Vowels are all underlyingly short.

b. Some geminates are lexically present.

c. The rhymes of stressed syllables must dominate two positions on the skeletal tier.

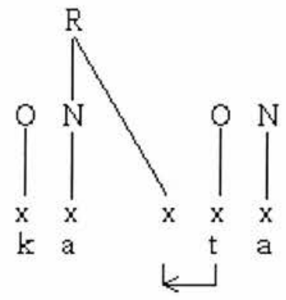
Geminate stops are easily accounted for by Gibb. A form such as 'kappi' [k^hahpɪ] (bully) is represented as in (22). The stop here is lexically geminate, and as such the BRC is automatically complied with as the rhyme already branches. No skeletal slot need therefore be added. This leaves open the question of what form the geminate consonant takes. Normally we might expect that the gemination in (22) would involve all the elements of the stop to occupy both coda and onset, as it does with the nonaspirates. However, aspirates choose only to spread the frication element h^o leftwards to occupy the empty rhymal slot. We must then make the further assumption that spreading is unidirectional, i.e. an element can spread in only one direction, so that if leftward spread is chosen in (22) h^o cannot subsequently also spread rightwards. This accounts for the absence of postaspiration in (22) and from preaspirates generally. Preaspiration i.e. [h] then consists simply of the element h^o spreading from an onset stop to a preceding empty rhymal slot, linking the release and aspirated properties of the stop directly with preaspiration's realisation as [h]. This of course avoids the problem noted in chapter two, namely the difficulty of equating the role of the larynx in creating voicelessness and postaspiration with the realisation of preaspiration as [h].

(22)



We can compare (22) with (23) for 'gata' [ka:tʰa] (gate) below, where the intervocalic stop is not lexically geminate and the vowel is assumed to be underlyingly short. The BRC requires the stressed rhyme to dominate two skeletal points, and as only one point is underlyingly present another is inserted. Geminate stops can occupy both the onset and coda positions (they both license the occurence of a coda and fill the coda position) and no further operations are necessary to allow the stop in e.g. kappi to occupy the coda. The BRC introduces a rhymal slot, and we would assume that in 'gata' the following onset will then govern that slot and spread h° into it, an unwanted result. To prevent this Gibb introduces the Cyclic Spreading Constraint (24). This prevents the /tʰ/ in 'gata' from spreading to create *[kahta]. Instead the vowel, the head of the rhyme, spreads rightwards to fill the empty slot.[17]

(23)



(24)

Cyclic Spreading Constraint (CSC)[18]

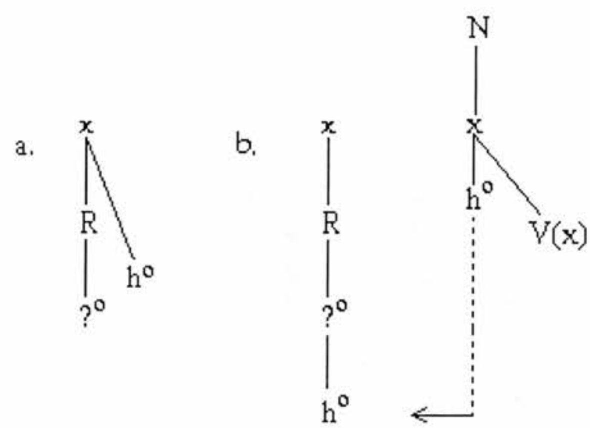
Non-nuclear segments cannot spread to points they have not occupied lexically within a cyclic domain

[17]We should note therefore that the creation of the rhymal point and the subsequent spreading of the nucleus are separate processes. Nuclear spread is a separate rule, one which operates as it were by default when nothing else is available to fill the slot.

[18]The use of the term 'cyclic' here is perhaps redundant given that GP does not subscribe to the phonological cycle as such. During the initial syllabification the aspirated stop in 'gata' is linked only to the onset. The BRC then creates a new domain, but the stop is prevented from spreading by the CSC. There is no ordering of levels implied.

Nevertheless, although the h° of the stop is prevented from moving leftwards by the CSC it must move somewhere, and Gibb suggests that in 'gata' it must move rightwards. Ignoring for the moment the question raised above as to the inadequacies of the internal structure of aspirates and why it is that h° should spread at all, we have two options as to the form of the spreading in 'gata'. The first is to 'break' the stop as in (25a), where the sequencing of the stop and noise portions of the stop is contained directly within the segment itself. The alternative is the form in (25b) where the h° spreads instead to the directly following point which is here the nucleus, forming a light diphthong with the nuclear vowel.[19]

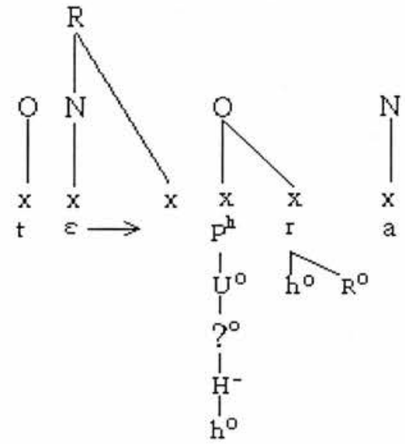
(25)



Support for the latter comes in the behaviour of aspirated stop + /j, v, r/ clusters. (26) shows Gibb's representation of 'depra' [tɛ:pr̥a] (sadness). These are interpreted as onset clusters, with the stops as head of the cluster. Again the nucleus of the stressed syllable spreads rightwards to occupy the vacant rhymal slot as the stop is prevented from spreading by the CSC. The h° of the stop must still spread and moves rightwards again to occupy the directly following skeletal point, though now this slot is both consonantal and in the onset, and instead of forming a light diphthong it forms a contour segment of some kind.

[19]The diagrams are somewhat confusing. In (25b) h° is initially only present in the stop. The arrow here indicates the slot *from which h° spreads* rather than the direction in which it spreads, as we might expect. In (26) below, Gibb does not show the spreading of the h° from the stop to the sonorant.

(26)

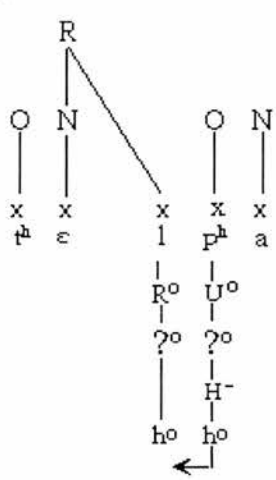


There are some obvious problems with this analysis. Nowhere else in the language is h^o, or any other consonantal element, able to occupy a nuclear, as opposed to rhymal, slot. Gibb argues that this is why postaspiration is shorter than preaspiration, which does not share a skeletal slot but instead occupies one of its own. However, the same contrast would be made just as simply if postaspiration shared a slot with the preceding stop of which it is a part. Further, this also implies that the vowel itself should be significantly shorter in this environment than when following other consonants which do not share its timing slot, perhaps only half its normal duration, but I know of no evidence to support this. We might also ask what other elements we can expect to spread in this way, either in Icelandic or in other languages, and what would be the effect of forming light diphthongs with e.g. $?^o$ or H^-?

The problem seems to stem from the spreading of h^o to following consonants. The simplest analysis of spreading would be to spread h^o directly to the skeletal slot of the following vowel, avoiding the problems noted above (but also implying either a fully voiceless vowel or a fricative), but if we were to apply this to following /j, v, r/ we would expect not simply voicelessness to result but frication. For example, were we to spread h^o to R^o, the resulting segment would be h^o, R^o, i.e. [s], not [r̥].[20] The only way to avoid this is to claim that a contour segment results, but even this is phonetically inaccurate, as we find [r̥] and not *[hr]. If we insist on a contour segment in this environment we must insist on a contour segment when the onset does not branch, i.e. (25). This is nevertheless an improvement over FG analyses using [+SG] or [Asp] where the phonological and phonetic ordering are either ignored or misrepresented, but the evidence seems to suggest that it is within the stop itself that ordering takes place. To claim instead a direct link between h^o and a following sonorant or vowel does not seem warranted by the facts.

Ordering is unarguably present, however, in sequences of liquids and nasals plus stops. (27) shows Gibb's analysis of 'telpa' [tɛl̥pa] (girl). The stop licenses the coda /l/ and h^o spreads in the direction of government, that is leftwards, to the coda, again causing the affected coda consonants to be realised as voiceless, though apparently without the necessity to create contour segments.[21] It is the presence of a governing relationship between the aspirated stop and a preceding coda position which allows h^o to spread, triggering preaspiration. Devoicing of coda consonants and loss of aspiration from the following stop are thus seen to be merely manifestations of h^o spreading and identical in form to the preaspiration seen in 'kappi'. This connection between voieclessness and aspiration is a necessary one.

(27)

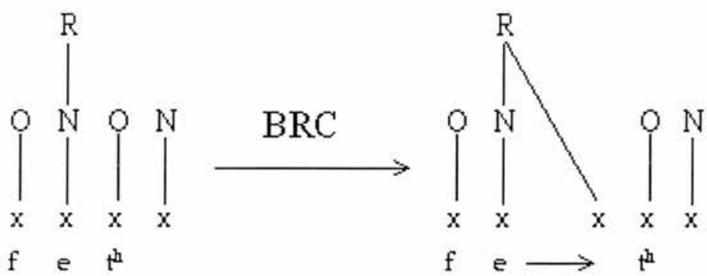


[20]This assumes that h^o is here the head, which may not be the case. However, if h^o is not the head then the problems of interpretation are simply made worse.

[21]There are similar problems regarding the phonetic interpretation of spreading h^o to coda consonants. For greater discussion of internal segment structure in GP see Harris (1990).

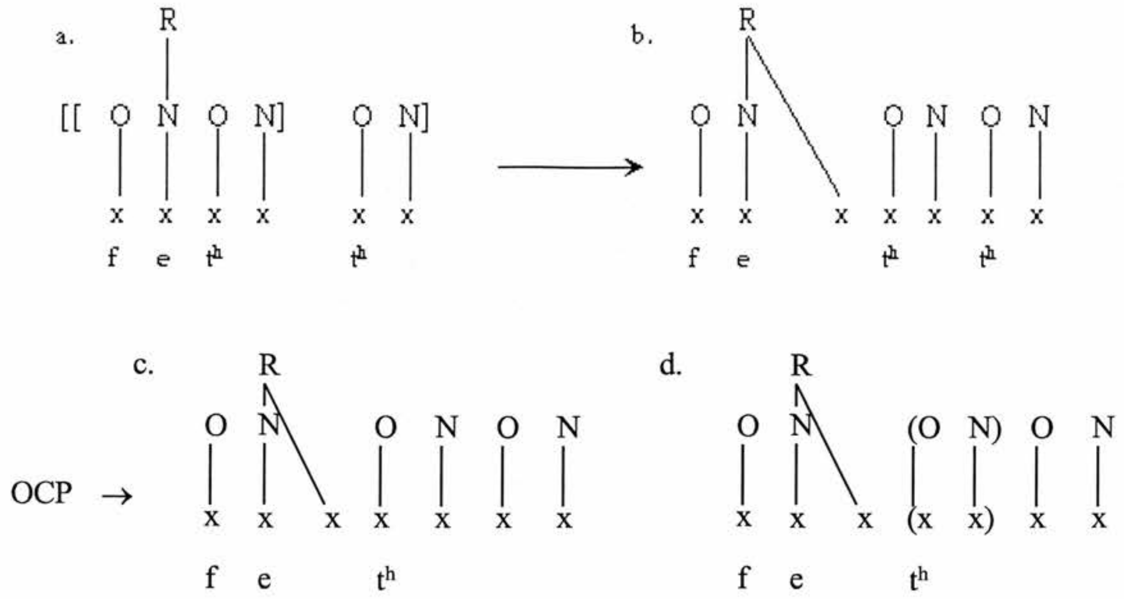
The combination of the BRC and the CSC account straightforwardly for the mutual exclusiveness in stressed syllables of long vowels in open syllables and short vowels in closed syllables. Preaspiration and long vowels and postaspiration and short vowels are also now clearly mutually exclusive, with both preaspiration and devoicing of coda consonants being reduced to the manner in which a coda rhymal slot is governed. Problems arise again, however, once we turn away from lexical to derived environments. (28) shows the derivation of the long nucleus in [fei:t^h] 'fat' (adj.). The BRC operates to create the available skeletal point to which the /t^h/ is restrained from spreading by the CSC, allowing the vowel to spread in its place.

(28)



In (29) we see the neuter form of the same adjective, formed by addition of the neuter suffix /t^h/. (29b) shows the immediate effect of the BRC in creating the rhymal slot, and again the stem final consonant is prevented from occupying the slot by the CSC. However, the vowel does not itself now spread as we might expect. Gibb claims that the addition of the suffixal /t^h/ sets up an OCP violation which is resolved by deleting one of the stops to leave only a single segment. (29c) clearly suggests that it is the suffixal /t^h/ which is deleted, but Gibb claims that it is in fact the stem final consonant which deletes. As the suffixal /t^h/ is not in the same cycle as the deleted stop it is free to spread to the empty rhymal slot, which Gibb represents as in (29d).

(29)



There are a number of problems here. It is unclear whether the bracketing in (29d) is meant to represent anything other than that the stop is in another cycle. In any case, if it is the stem consonant which deletes, presumably the resulting sequence of empty onset and empty nucleus must also be deleted, that is it is not only the segmental content which is deleted as a result of the OCP violation but the entire skeletal structure. This is neither made clear in Gibb's representation, nor is it clear why it should be so. Assuming however that the suffix is now adjacent to the created rhymal slot, Gibb suggests that it is able to spread to this point as the CSC does not apply given that the suffix is in a separate cycle. This assumes that the suffix is now able to license the existence of the rhymal slot despite the lack of any preexisting relationship between the two.

However, there is no indication as to what process prevents the nuclear vowel from spreading to the vacant point in the first cycle, bleeding the spreading from the suffix. Spreading here must somehow be blocked, and blocked in general until all possibilities for spreading from a following stop have been exhausted, including spreading from stops in different cycles. In addition, it now seems clear that rather than adding a simple skeletal point with no higher structure, the BRC must add specifically a coda slot. GP does not distinguish at the segmental level between glides and vowels, preferring to distinguish between them at the syllabic level, so that e.g. I° is a vowel /i/ when associated to a nuclear slot, but a glide /j/ when it is associated to an onset or coda. As there is no process specified to convert the BRC-created point into one dominated directly by the nucleus rather than by the rhyme when government by a following onset is not possible, once the vocalic elements spread to this consonantal position the resulting rhyme should contain not a long vowel but a sequence of short vowel plus semivowel.

In order for the OCP to apply, two segments must be adjacent at some level of the representation. For this condition to be met in GP there must be some kind of governing relation between the two segments, so that in 'feitt' one stop must somehow govern the other. The empty nucleus of the stem is domain-final in the first cycle and is therefore licensed automatically, so there is no reason for the two stops to show any governing relationship.[22] Without such a relationship it is difficult to see how the OCP can apply.[23] Further problems arise when the neuter suffix is added to words with final consonants other than /tʰ/, such as 'heilt' [heil̥t] (whole) as in (30). There is no obvious OCP violation here and hence no deletion of the stem-final liquid.[24] If we assume that there is some kind of governing relationship between the stops in (29) we would obviously assume the same kind of relationship to be present here also. We might therefore expect *[hei:ltʰ] or perhaps *[hei:l̥t] as the resulting forms, with a long vowel followed by either a voiced lateral and postaspirated stop or a voiceless lateral and unaspirated stop, depending on whether or not h° spreads. The /l/, being in the same cycle as the preceding vowel, is unable to spread, and the suffix has now no rhymal point to which it can spread.

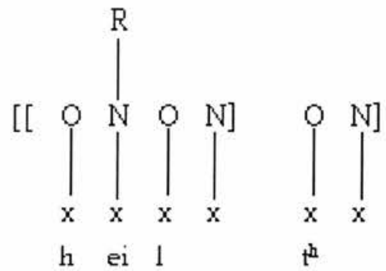
The same can be seen in 'rakt' [raxt] (moist) in (31a). Gibb represents 'rakt' orthographically as 'raxt' as if the 'x' were underlying and gives the representation in

---

[22]If this were not the case then we would expect the nucleus to remain unlicensed and thus give rise to an apparent epenthetic vowel.
[23]We shall also see that given Gibb's assumptions as to the direction of onset to onset government in Icelandic, it is the stem-final stop which would be the head. We must then explain why it is that the head is deleted to avoid the OCP violation and not the non-head.

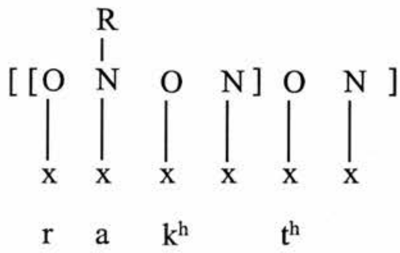
(31b). In fact this is not a possible spelling - 'rakt' is the neuter form of which the feminine is 'rak' [ra:kʰ]. The form [raxt] is the result of a phonological rule of spirantisation with which I shall deal in greater detail below and in chapter 5. Again if the suffix is in a separate cycle there is nothing which would suggest either that the stop spirantise or that it license the preceding rhymal slot and then spread. Instead we would expect the nuclear vowel itself to spread and result in *[ra:kʰtʰ] instead of [raxt].

(30)

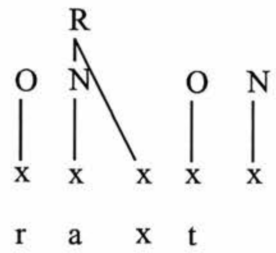


(31)

a)



b)



The CSC then is not able to fully explain why spreading of h° is blocked in forms such as 'feit' but not in 'telpa' or 'feitt'. This results from confusion as to the aim of the CSC. Aspiration occurs in 'kappi' because the stop lexically occupies and governs the rhymal slot. It occurs in 'telpa' again because it governs the coda. In 'gata' the BRC applies and we expect preaspiration as the medial stop should govern the created coda. As it does not the CSC is created to block spreading. This same CSC however prevents the correct derivation of 'heilt' and creates the difficulties in the analysis of 'feitt'. The general principle of aspiration following the path of

---

[24]We cannot claim any real OCP vioation due to the coronal nature of both the sonorant and the stop, as non-coronals occur in the same environment.

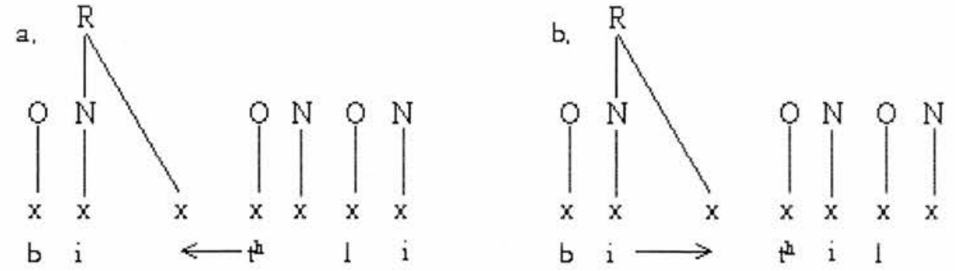
government, however, seems a sound one. GP's principal difficulty lies in its conception of the rhyme. In order for the stop to govern and spread h^o the created slot must be a coda, but in order for government to be blocked and for the vowel to be long the created slot must be nucleic. There seems no way around this paradox.

The vital role of both government and empty nuclei in preaspiration is shown clearly in Gibb's analysis of stop + /l, n, m/ clusters where preaspiration occurs without the presence of geminates. Although, as noted above, these sequences are generally illicit as word-initial clusters they do occur within the word, both in lexical and derived contexts, and in all cases result in preaspiration. In words such as 'epli' there appears to be an underlying cluster of some kind, while in 'Bitli' we can derive the cluster by a regular process of epenthesis. In either case the surface environment is identical.

Gibb represents 'Bitli' as in (32a). /t + l/ is ruled out as a coda-onset sequence by a set of principles relating to segmental strength. Put simply, /l/ is 'weaker' than /t/ and thus unable to govern it. This leaves (32a) as the only possible syllabification.[25] We can compare this with 'Bitil' in (32b) where syncope has not occurred. Some process is clearly at work which both deletes the stem-final vowel and allows the aspirated stop to govern the metrically created rhymal node, a process triggered by the addition of the dative singular and other suffixes ((10) above). The crucial point is that the stop be able to govern leftwards. Without this government there can be no connection between preaspiration in 'Bitli' and in 'kappi'. We must also account here for forms such as 'vopn' where there appear to be not one but two empty nuclei.

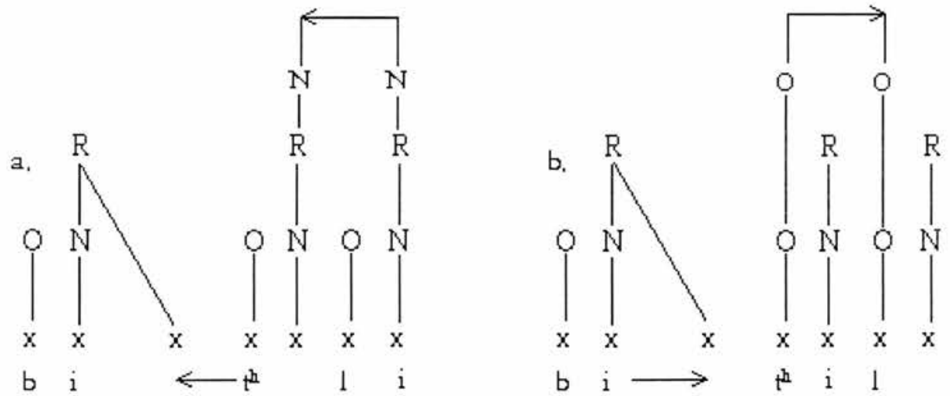
[25]Gibb attempts to rule out /tl/ as a possible onset by claiming that it would create an OCP violation. She further claims that this is a universally illicit onset. This is contradicted by languages such as Gaelic where both /t^hl/ and /tl/ are licit clusters word-intially. It seems simpler to simply state that such clusters are illicit in Modern Icelandic, though they were not so in earlier stages of the language.

(32)



Gibb suggests that the syncope is the result of some kind of governing relationship between the aspirated stop and the following onset's liquid or nasal at the interconstituent level. However, Gibb seems to over-account for this goverment, as it were. In order for the two onsets to see each other there must be no intervening segments at their level of projection. In other words, the two onsets must be adjacent at some level. In 'Bitil' a full vowel intervenes - the final nucleus is an empty syllable and thus unable to govern the preceding nucleus, which is then realised. This nucleus then projects onto the nuclear projection level (33b), blocking the two onsets from seeing each other, and preventing the /tʰ/ from governing the /l/. The addition of the dative suffix in (33a) provides the necessary governor for the preceding nucleus. Gibb assumes that the dative /i/ is able to govern the stem-final vowel and that as a properly governed nucleus this vowel becomes a licensed empty category. As licensed empty categories are not projected the two onsets are now visible to each other and the /tʰ/ is able to govern the /l/.

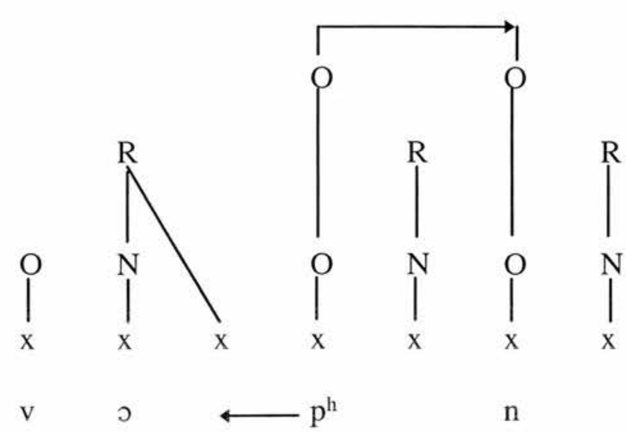
(33)



The conversion of the stem-final /i/ of 'Bitil' into an empty vowel simply because it is licensed seems at odds with the basic principles of GP. Government of full vowels does not normally result in such a change. In fact, it seems as if it is the relationship between the two onsets which is important here, and it would be easier to claim that in this environment an onset to onset governing domain is set up regardless of the content or lack of it of any intervening nuclei. Such a process is needed in any case for words such as 'vopn' (34) where it is only the onset to onset governing domain which allows the intervening nucleus to be governed and thus licensed, as there is no following full vowel available to govern it. In addition, the stem-final vowel in (33a) has to be projected in order for it to be governed in order for it not to be projected. This seems at best paradoxical and at worst it would require some kind of cyclic interpretation of interconstituent government.

Regardless of these problems, the crucial point is that the /t^h/ in 'Bitli' now governs the following onset and is a governor at the interconstituent level. This government filters down, so that a segment which has governing status at any level of a derivation will have that status at all other levels. This being so, the /t^h/ is now able to govern the BRC created slot and preaspirates to give [bɪhtlɪ]. This raises questions again as to the nature and the purpose of the CSC. As formulated it still does not appear to matter whether or not the /t^h/ of 'Bitli' can govern leftwards or not, as it is not government that is outlawed but spreading.

(34)



More fundamental problems arise from the very notion of government and its physical realisation. The direction of interconstituent government in GP is generally assumed to be right to left (Kaye 1990) and Gibb follows this in nucleus to nucleus government. While government within constituents is constrained by the strength ('charm') of segments as already noted, no such constraints are present between nuclei. Gibb however assumes that the direction of government from onset to onset in Icelandic is left to right, at least in the syncope environments crucial to preaspiration. While there is no discussion as to the crosslinguistic consequences of such a change, it nevertheless seems justified to the extent that it works. In 'Bitli' the /l/ is not strong enough to govern the /tʰ/ within an onset, ruling out 'lt' as a possible onset cluster in any language, while 'tl' is licit, though rare. Gibb carries this pattern over into interconstituent government which in 'Bitli' runs left to right, /tʰ/ governing /l/.

Direction of government then differs according to whether it is between nuclei or onsets. Further, theoretically at least, syncope could be found in which the rightmost consonant were stronger than the preceding, and government might run here from right to left. As long as the leftmost consonant is stronger, however, government should always be left to right. Unfortunately this is contradicted by syncope in other environments in Icelandic. While Gibb notes that syncope in Icelandic always involves a stop followed by a liquid or nasal, she does not note the different outcomes when that liquid is /r/ and not /l/. In 'akur' [a:kʏr] (field) we find the expected long vowel in the stressed nucleus. In the dative singular 'akri' [a:kr̥ɪ] and nominative plural 'akrar' [a:kr̥ar̥] we find syncope of the same kind as in 'Bitli' (the final /r/ is here part of the stem, not of the suffix) in precisely the same environment, yet rather than preaspiration we find a long vowel followed by a postaspirated stop. Government must still run from left to right as 'rk' is not a licit onset while 'kr' is not only licit within GP but is licit within Icelandic. Thus it seems as if 'akri' must have the same onset to onset government as in e.g. 'Bitli', yet if this is so then GP fails to predict the correct surface form. The pattern is reversed in 'hamrar' [hamrar̥] nominative plural of 'hamar' [ha:mar] (hammer), where after syncope the vowel is shortened as it is in 'Bitli'. Although this is not a preaspiration environment it is otherwise identical. The /m/ is in addition lengthened and can be regarded as geminate and this lengthening is found throughout the language in this environment (Hermans 1984). In other words, it seems as if the /m/ behaves just as the /tʰ/ in 'Bitli', governing the preceding rhymal slot and spreading itself to occupy it.

Gibb's analysis fails to provide an explanation for these facts. There is no explanation possible within GP as to why the identity of the liquid should produce such different surface patterns. Nor is there any explanation for the lengthening of /m/ in 'hamrar' as the spreading from /tʰ/ is specifically attributed to the desire to spread h° and not to any general theory of onset-coda government. Gibb's analysis relies on the creation of a basically consonantal metrical slot in the stressed rhyme, a slot which is governed by a following onset head. This onset head must be prevented from governing the slot in a large number of cases, and allowed to do so in others, necessitating the introduction of the CSC. I shall attempt to show below however that an analysis employing the same basic conceptions of syllable structure is both possible and desirable within AP.

4.3.4
Icelandic Syllable Structure in AP

We can immediately discount any explanation of the Icelandic data in terms of changes in overlap or gestural stiffness. Altering the stiffness of vocalic and consonantal gestures, until we achieved the required length, would not account for the complex pattern of changes we have observed, and it is doubtful in any case whether such large and discrete differences could be effectively handled by such a change. Accounting for vowel duration in terms of increasing or decreasing overlap between vowels and consonants is in any case immediately ruled out by forms such as 'á' [au:] (river), where there is simply no consonant, either preceding or following, which could be shifted to produce the necessary result. There is no reason, however, why AP should not employ basically the same mechanisms as the other theories discussed here (though I am not, of course, advocating the use of timing units such as syllable slots or moras). In other words, where a short vowel consists of a single unit of some kind, so a long vowel consists of two of those units. The question to be then asked is what the identity of such units might be.

As we saw in section 1, most approaches to syllable structure assume that binarity is a fundamental characteristic, and I assume that the same should be true of a gestural approach. Browman & Goldstein (1986) touch upon the possibility of such an approach in their discussion of the nasals and prenasalised stops in kiChaga (35). Although both the nasals and the prenasalised voiced stops seem to be single segments, the voiceless prenasals are realised as syllabic nasals followed by homorganic stops, resulting in an overall longer segment or cluster. By comparing the labial gestures of the inital segments in (35) Browman & Goldstein noted that the gesture in /mpaka/ was clearly longer than that in any of the other words, including the prenasal in /mbaka/. The difference in duration, moreover, was concentrated in the central 'holding' portion of the labial gesture, the onset and offset slopes being almost identical for all words. This identity rules out a process which would lengthen the gesture as a whole, as this would act to make these slopes less steep. Some process is then obviously at work which holds for the peak of the gesture alone, causing it to lengthen. Rather than introduce some new force which would account for this, they note that the correct result can be achieved if we assume that instead of a single longer labial gesture, the voiceless prenasal in /mpaka/ may consist of two separate overlapping gestures.

(35)

/wia mboka _____ kinbuho/ 'say to the starter _____ slowly'

a. paka	'cat'
b. maka	'year'
c. mpaka	'boundary'
d. mbaka	'curse'

Recalling the structures described in chapter 3, I proposed that in order to adequately capture a number of phonological generalisations, segments should be described in terms of gestures forming abstract head-dependent structures. For example, a segment such as /t/ would contain a tongue tip gesture which would act as the head, along with a non-headed glottal gesture. In essence, the gestures together with their phonological relationship form /t/. While the discussion concentrated on the structures of consonants, I assume that the same types of processes also create vowels. Thus, if a vowel such as /a/ indeed consists of a single TB : pharyngeal : wide gesture, then if every segment must contain a head there is clearly only one candidate for that head, namely the TB gesture. Without the phonological structure the TB gesture is simply that, a TB gesture, and not a phonological object.

In common with other theories, and in line with the claims of Browman & Goldstein (1989), I assume that vowels are the focal point of syllable structure, and in addition that they can be fully described in terms of the structures set out in chapter 3. Taking for our initial example a word which consists solely of a vowel, e.g. /a/, I propose that the vowel is the head of its domain, that domain consisting solely of the vowel itself. We can represent this simply as V, where V represents any vowel segment i.e. a segment which obligatorily contains at least one headed gesture, and optionally one or more headed or non-headed gestures. The V does not associate to a higher level node of any kind, no timing slot, no nucleus or rhyme. V forms a constituent by virtue of its status as a head without having to associate to any preexisting syllable node, echoing the structures of Dependency Phonology (Anderson & Ewen 1987) rather than those of GP, and we can refer to this constituent as the nucleus.

(36a) is a graphic representation of a word consisting of a single headed vowel such as /a/, where the V is a head. As in GP, all heads can license a dependent position, so that in (36b) the head licenses the presence of a following dependent vowel. The existence of the dependent vowel is therefore licensed by the head in a way comparable to the autosegmental licensing of Goldsmith (1990), so that in the appropriate circumstances a headed vowel has the ability to create, as it were, a separate though dependent vowel. Adapting the term of Roca (1994) for GP, we can refer to this as *direct licensing*, where the existence of V^2 in (36b) is dependent on the prior existence of V^1. A segment is directly licensed if it enters a *direct government* relationship with another segment. In (36b) V^1 *directly governs* V^2 where direct government is a dependency relation which is binary, intransitive, asymmetric and irreflexive so that for any sequence *x,y* where *x* directly licenses *y*, *x* directly governs *y*. As in GP, the head of a domain does not require to be directly licensed within that domain, so that in (36b) V^2 is directly licensed while V^1 is not. The solid arrow notation in (36b) represents both direct licensing and direct government.

(36)

a) V^1 b) $V^1 \longrightarrow V^2$

c) $V^1 \dashrightarrow V^2$

Direct government is thus a dependency relation of the same type as in DP and in addition encodes constituent status in precisely the same way. In (36a,b) we can recognise two constituents, each corresponding to the traditional notion of nucleus, where a constituent is defined as any segment and all those segments subordinate to it, subordination being the transitive closure of dependency (Anderson & Ewen 1987). In (36a) this gives us a constituent consisting of a single headed vowel, while in (36b) the constituent consists of a headed vowel and its dependant. In traditional terms we can refer to these as non-branching and branching nuclei respectively. This again follows the notation of DP, where each category is immediately associated with a terminal, and where the segments themselves are the syllable positions, doing away with the need for any superordinate nodes. Further, the dependency relation means that heads are obligatory, so that e.g. a nucleus cannot exist without there being a nuclear head.

Let us assume that (36b) represents the diphthong /au/, where /a/ is the head which directly licenses and directly governs the non-head or dependent /u/. The relationship between the two segments, /a/ as head governing the non-head and dependent /u/, captures the constituent structure, but it tell us nothing, however, about the physical relationships which exists between the two segments. In other words, direct licensing and direct government make no claims of any kind as to the physical coordination between them i.e. direct government does not entail coordination.[26] Given the physical nature of gestures it is essential that we account for this coordination, as if V^2 were to fail to coordinate with V^1 then the resulting structure would not be that of a diphthong but a hiatus sequence /a.u/. Thus, in order for the dependent V^2 to be fully integrated into the prosodic hierarchy it must coordinate with V^1.

This integration is achieved through *indirect licensing*. Every segment within a word must be indirectly licensed, where a segment is indirectly licensed if it enters an *indirect government* relationship with another segment. Indirect government, which like direct government is intransitive, asymmetric and irreflexive, follows the direction of direct government so that if x directly governs y and an indirect government relation also holds between x and y, then x indirectly governs y. Note that if x directly governs y, this does not automatically mean that x must also indirectly govern y, as although this generally holds true we shall see circumstances in chapter 5 where direct government occurs without indirect government also occurring. Further, indirect government may hold between two segments which do not also show direct government, so that while direct government often implies indirect government, the reverse is not true. The most important distinction, however, between direct and indirect government is that the latter entails physical coordination, i.e. if an indirect governing relation exists between x and y, then x and y coordinate.

Thus, in (36b) where V^1 directly governs V^2, V^1 also indirectly governs V^2 and the two coordinate. Indirect licensing and government are shown by the dotted arrow notation in (36c). Direct licensing therefore sanctions the existence of V^2 while indirect licensing integrates it into the prosodic hierarchy. A problem may appear to arise with (36a) in that when a word consists of only a single vowel, there

[26] We shall see below that physical coordination between segments involves *indirect* government.

is nothing else present within the word with which the vowel can form an indirect governing relationship and therefore (36a) would appear to violate the requirement that every segment be indirectly licensed. However, it is clear that as indirect licensing acts to incorporate segments into the prosodic hierarchy, there is no requirement for the vowel in (36a) to be indirectly licensed as there no other segments present in the word other than the vowel itself. I stop short of stating that vowels which are the head of their domain do not require to be indirectly licensed, because as we shall see below in the discussion of Italian and Turkish, there are languages in which this appears not to be the case.

How, though, can we interpret the structure in (36b) as representing a long vowel? A short vowel /a/ consists of a single segment formed from a single gesture, as in (36a). In contrast, the long vowel /a:/ consists of two segments, but crucially I suggest that this is also formed from a single gesture. In other words, just as the head vowel in (36b) directly licenses and governs the dependent vowel, a long vowel is simply one in which the head also demands that the gestural content of each vowel be identical. The result of this is that for /a:/ the TB : pharyngeal gesture is effectively realised twice, a single gesture being required to form two separate segments. The single constituent nature of the two segments is reflected in their coordination: the head and dependent vocalic segments in the nucleus will coordinate the offset of the head vowel with the target of the second to form an overall longer nucleus. This has the effect as it were of coordinating the gesture's offset with its own target, so that just as the initial gesture is turned off the second reaches its target, resulting in a long vowel. The TB gesture for the vowel /a:/ is thus marked as filling the position of both head of the syllable and dependant, acting as both a single unit and a cluster of two vowels. There is no need to lengthen the vowel in any way by changing the gesture's stiffness or by a change in coordination or overlap.

Thus we make basically the same claim regarding vowel length as other non-gestural theories such as FG, DP or GP, namely that short vowels consist of one of something and long vowels of two of something. Informally we can view the stressed vowels of Icelandic as containing a branching nucleus, though this view of nucleus branching differs from that of GP and FG in a number of ways, and has a number of consequences for syllable structure in general. For a vowel in GP we find a nucleus node governing a timing slot governing (or perhaps just linking to) segmental material. In AP on the other hand the head vowel simply licenses the appearance of another vocalic segment which already has its own internal duration and segmental structure, creating a simpler structure which is still binary but without the need for external timing units of any kind.

As discussed earlier, vowels are viewed in AP as being overlaid by consonantal gestures, and the patterns of coordination suggested by Browman & Goldstein have reflected this directly, creating structures in which vowels and consonants are on separate planes in a way reminiscent of McCarthy (1979). One effect of this is that vowels and consonants cannot occupy the same plane, effectively ruling out the existence of the rhyme in AP, as I shall discuss below. Licensing relationships do hold, however, between vowels and consonants, and I suggest that a vocalic head can directly license and govern a dependent consonant in the same way in which it directly licenses and governs a dependent vowel (37a). V^1 and V^2 together form a single constituent, and similarly V^1 and C (or V^1, V^2 and C) also form a constituent. The direction of government is again shown by the arrow, and the consonant is placed below the vowels, reflecting the fact that each are in different planes. The consonant and dependent vowel in (37a) are distinguished in one further way, in that the consonant, although governed by V^1, is licensed as the head of its domain, and as a head it can directly license and govern a dependent consonant C^2 as in (37b). While C^2 is not directly dependent on V^1, it is subordinate to it following the chain of dependency from V^1 through C^1 to C^2, although V^1 and C^2 do not together form a constituent.

(37)

a)
$$\begin{tikzcd} V^1 \arrow[r] \arrow[d] & V^2 \\ C^1 & \end{tikzcd}$$

b)
$$\begin{tikzcd} V^1 \arrow[r] \arrow[d] & V^2 \\ C^1 \arrow[r] & C^2 \end{tikzcd}$$

c)
$$\begin{tikzcd} V^1 \arrow[r] \arrow[d, dotted] \arrow[dr, dotted] & V^2 \\ C^1 \arrow[r, dotted] & C^2 \end{tikzcd}$$

Again, however, it is not sufficient to say that the various consonants are directly licensed and governed, they must also be indirectly licensed and governed. As noted above, Browman & Goldstein (1990b) identified the C-centre as a point in the vowel around which onset consonants clustered, the consonants themselves showing greater than expected overlap, and we need to account for this fact. The consonants in (37b) form a single constituent through direct government in the same way as do the vowels, and like the vowels I suggest that the dependent consonant is indirectly licensed by being indirectly governed by the head. As the relationship between branching onset consonants is of the same type as that between the vowels in a branching nucleus, so too is the coordination, so that the offset of C^1 is coordinated with the target of C^2. Were the consonants to share a different and less close phonological relationship, such as that between coda and onset, the coordination would differ, as we shall see below. This new 'macrosegment', created by the government of C^2 by C^1, then appears to coordinate as a single constituent with the C-centre of the vowel. This suggests that the dependent consonant is indirectly licensed and governed not only by onset head but also by the head vowel, and this is represented in (37c) by the dotted arrow between the two.

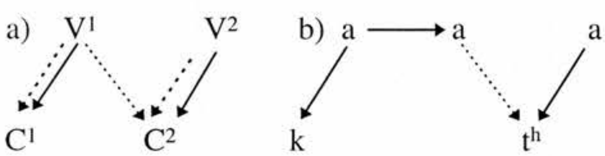
In fact, the notion of a C-centre as such is now removed, as its main effect, that of altering the coordination of the consonantal gestures, is now a result of the phonological relationships between the consonants. However we can retain the use of the term C-centre to refer to the fixed point early in the cycle of the vowel to which the consonants coordinate. Should it be the case that the C-centre is found to be non-universal, then this would suggest that for those languages the dependent C^2 in (37c) is not indirectly licensed by the vowel. What remains true for all of these languages, however, is that the structures in (37) entail recognition of both onsets and nuclei as constituents. Nevertheless, assuming the universality of the C-centre leads us to make the further claim that all consonants must be indirectly licensed by a vowel, whether or not they are also directly licensed by a consonant. We shall examine the consequences of such a claim below, and more closely in chapter 5.

The structures in (37) are enough to describe monosyllabic C(C)V(V) syllables, but are inadequate on their own to describe words such as Icelandic 'gata' [ka:ta] (gate). Although syllable structure is built up from abstract phonological objects, with gestures forming abstract phonological relationships to create segments, such gestures are nevertheless real physical objects. While the structures proposed so far could describe the two syllables [ka:] and [t^ha] as separate entities, the physical nature of gestures means that within a word these two syllables must physically coordinate, and further this coordination must be such that the vowels of each syllable overlap. We know from observation that some coordination is present between the two syllables, but two questions remain to be answered. Firstly, is the overlap between the vowels direct or indirect, in other words do the vowels of [ka:] and [t^ha] coordinate directly with each other or only indirectly through their coordination with the consonants? Secondly, whatever type of coordination is proposed, what are the phonological consequences?

For 'gata', as [k] is word initial and directly and indirectly licensed by the following vowel, there are no additional coordinative requirements other than that it coordinate with that vowel. The same is not true of [t^h] however. While [t^h] is indirectly licensed by, and hence coordinated with, the following [a], in order for it to be fully integrated into the prosodic hierarchy, [t^h] must also be coordinated with the immediately preceding syllable. The target of the consonant therefore forms an indirect government relationship with the preceding nucleus, coordinating its target with the offset of the preceding vowel and the onset (or C-centre) of the following, with the result that the two vowels overlap without being directly coordinated.

This results in the structures in (38). For a CV$CV word (not a possible form for Icelandic), coordination is as in (38a). As onsets, both consonants are directly licensed by their following vowels, as well as being indirectly licensed by them, this dual licensing being indicated by both the solid and dotted arrow notation (in further diagrams, this double arrow notation will be represented by the solid arrow alone, except where this may cause confusion). In addition C^2 is also coordinated with the preceding vowel, shown here by the dotted line. In other words as C^2 is not word-initial, in order for it to be fully integrated into the prosodic hierarchy it is indirectly licensed by both the following and the preceding vowel, the direction of government again being shown by the arrow notation. (38b) shows the structure appropriate for 'gata' with a CVV\$CV structure. Again the onset of the second syllable is coordinated with the immediately preceding vowel, but because the preceding vowel is long C^2 coordinates not with the head of the preceding syllable but with its dependant. This coordination and concomitant indirect government has an important role in the structure of syncope and preaspiration, as discussed below. For the moment let us simply note that what is important is not that C^2 enter an indirect licensing relationship with the immediately preceding vowel as such, but that it enter an indirect licensing relationship with the preceding syllable in some way.

(38)



We can now describe strings of CV(V)CV syllables, and this may be adequate for some languages with only open syllables. How, though, do we account both for the existence of closed syllables and for the fact that long vowels and short vowels in closed syllables are somehow identical in terms of syllable weight in some languages? The simplest way to capture the fundamental identity between the stressed syllables in 'gata' and 'telpa' (girl), without referring to abstract timing units of any kind, is to claim that despite the difference in surface duration the vowels in each syllable are in fact affected in the same way by the BRC (or whatever we term the process), i.e. both syllables branch and can be regarded as containing long vowels. While this might appear to be false if we simply compare the surface length of the stressed vowels in each, the general principles of coordination of AP demand that this be true.

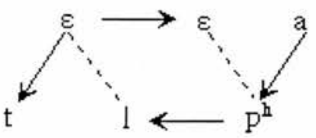
Let us assume that we have a word of the form C^1VC^2 \$ C^3V, where C^2 occupies a coda. In order for VC^2 to form a constituent comparable to a long, heavy vowel, the vowel would have to directly license the consonant, just as it directly licenses an onset consonant. The two would be distinguished however by the fact that the coda consonant would not be a head (assuming that VC forms a constituent parallel to VV) and would have to be on the same plane as the vowel. We would then need to explore the consequences of consonants appearing on both the vowel and consonant plane, and in addition we would be left without a simple explanation of phenomena such as compensatory lengthening following coda loss. Alternatively we could follow the reasoning above and assume that VV and VC are equivalent in weight because VC somehow contains an underlying long vowel. How could this be achieved?

The AP equivalent of Gibb's BRC is the Branching Nucleus Constraint (BNC), which simply states that if a headed vowel is stressed it must directly license and govern a dependent vowel. We can regard the BNC as applying across the board, regardless of the identity of any following consonants, to lengthen the stressed nucleus. Taking as our example 'telpa', the BNC will act upon the initial, stressed vowel, thereby creating an initial syllable [t^hɛ:], leaving aside for the moment the relationship between this sequence and the /l/. The /p^h/ is the head of an onset, and we can thus expect it to coordinate with both the following and preceding nuclei in the way already established for onset heads i.e. as head of the onset it coordinates with the immediately preceding vowel. Thus, if were it not for the presence of the /l/ in 'telpa' the surface form would be [t^hɛ:p^ha] with a long vowel and postaspiration of the labial stop. It would be identical in structure to 'gata'.

The evidence from GP suggests that coda consonants are always dependent upon the prior existence of a following onset, and as we have seen this is reflected in the presence of a licensing relationship between the two to the effect that codas (or more precisely, rhymal adjuncts) are licensed through being governed by onsets. In a coda-branching onset sequence such as in English 'paltry' the coda /l/ must be part of a separate constituent to the onset /t/ as otherwise we would contravene the requirement that government be strictly directional. Similarly, in terms of the structures developed here, if the coda is a separate constituent rather than the rhymal adjunct of GP, then it requires both direct and indirect licensing. I suggest we follow GP in claiming that onsets directly license, and directly govern, codas so that the /l/ of 'telpa' is directly licensed and governed by the following /pʰ/. Unlike GP, however, I suggest that this /l/ does not form part of a rhymal adjunct but is instead in a separate constituent from either of the flanking nuclei or the following onset. One immediate consequence of this is that the /l/ must be a head and it should therefore show the general characteristics of heads, primarily the ability to license a dependent, non-headed segment within its domain. In other words, we predict the existence of branching codas. Such structures are explicitly ruled out in GP, despite the frequency of apparent branching codas in many languages e.g. English 'first', though I shall not explore this matter further here.

The /l t/ sequence of 'paltry' is thus distinct from the /t r/ sequence in the same word by virtue of the differing status of the two sonorants, a head of a coda constituent in the case of /l/ and an onset dependant in the case of /r/. The two are further distinguished by the manner in which they are directly licensed. Turning to 'telpa', it is clear that the /l/ coordinates with the following stop, and hence we can say that codas are both directly and indirectly licensed by forming indirect government relationships with following onsets, government in both cases running from right to left as in GP. Coordination between coda and onset in 'telpa' is distinct from that of a branching nucleus, simply offset of the coda consonant to onset of the onset consonant so that the two minimally overlap. We must recall, however, the earlier claim that all consonants must be indirectly licensed and governed by a vowel, and as indirect government is intransitive the /a/ of 'telpa' is unable to license the /l/. Hence, the coda /l/ must be indirectly licensed by the preceding vowel. The preceding syllable contains not one but two vowels with which a following consonant can coordinate, and an onset head, I suggest, licenses a preceding coda consonant to coordinate not with the immediately preceding vowel, the dependent vowel, but with the head of the nucleus. In other words, while the onset head /pʰ/ in 'telpa' is indirectly governed by and hence coordinates with the offset of the rightmost vocalic gesture, the /l/ is indirectly governed by and hence coordinated with the head of the preceding nucleus as with the /pʰ/ itself.[27] The full syllabification of ‘telpa’ is shown in (39).

(39)



There are a number of consequences of this interpretation of syllable structure. As in GP, codas are only possible if directly licensed by an onset. While the stressed nucleus of words such as ‘telpa’ remains long, its second half as it were is hidden by the overlapping consonant so that it appears on the surface to be short. The fact that both VV and VC syllables are interpreted as heavy syllables now reduces trivially to the fact that in both the vowel is long. Alternatively, rather than refer to the underlying structure, some languages may opt instead to base syllable weight on the audible length of the vowel and thus distinguish between the long open and short closed syllables. We also avoid the problems of distinguishing between rhymal and nuclear positions created by Gibb's analysis, as there is no equivalent of the rhyme, and thus no need for an AP equivalent of the Cyclic Spreading Constraint.

The confusion, noted by Brentari & Bosch (1990), as to whether consonants should be linked directly to syllable nodes or to moras does not arise, nor is there any need for abstract timing units of any kind. We also meet the demands of Browman & Goldstein (1992b) that ultimately all higher level units should be describable in terms of gestures and the relationships between them. The different types of gestural coordination noted by Browman & Goldstein (1988) are directly incorporated here. Onset consonants behave very much like segments with respect to their coordination

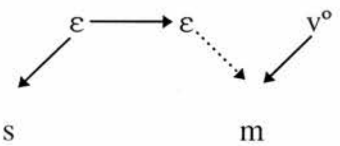
[27] This type of coordination seems to be true for Icelandic and most other languages. If on the other hand both the onset head and the coda consonant were to coordinate with the immediately preceding vowel we could expect to find forms such as CVVC$CV, with what might be described as trimoraic structures. Such structures, though rare, are found, e.g. Finnish (Dunn 1990) and require further investigation.

with a following vowel due to the phonological relationship between them and the following vowel, while coda and onset sequences do not show the same phonological relationship with the following vowel and thus do not behave physically in the same way. The same is again true of the relationships of both coda and onset with a preceding nucleus. Nevertheless, at the same time we are able to maintain the necessary vowel to vowel coordination across the intervening consonants as the presence of the coda consonant does not affect the coordination between the onset (or left edge of the C-centre) and the immediately preceding vowel. It is also clear that the coda consonant itself does not contribute to syllable weight, as both moraic theories and Browman & Goldstein claim, as this resides solely with the vocalic gesture, reflecting the central role of the nucleus in syllable structure. The mora now has no role to play in describing syllable weight and therefore no role to play in syllable structure per se.

The structure of words such as 'sem' [sɛ:m] (as, like) face the same initial problems of syllabification as they do in GP, namely that what would generally be regarded as a coda /m/ cannot be syllabified as such as there is no following onset which would either directly license it or govern it. The view of syllable structure being developed here implies that all consonants are ultimately dependent upon vowels for their existence and must therefore be licensed by them in some way. In terms of expected patterns of coordination, we have seen how the segments in both the onset and the nucleus show a substantial amount of overlap, reflecting the phonological relationships within them, and we would therefore expect a similar type of coordination between the nucleus 'sem' were it to directly license the following /m/ as a coda. Recalling the analysis of Browman & Goldstein (1990b) regarding final consonants, /m/ in 'sem' would on their interpretation not show this expected coordination but would show a far lesser amount of overlap with the immediately preceding vowel. In addition, as closed syllables in Icelandic contain only short vowels, phonologically the /m/ does not behave as if it closed the preceding syllable, i.e. it does not shorten the vowel. Instead it behaves in a parallel fashion to the intervocalic stop in 'gata', with neither consonant contributing to syllabic weight, despite the word-final /m/ being presumably realised largely within its own time period as Browman & Goldstein would claim.

If a final consonant does not form a constituent with a preceding vowel and is thus not directly licensed by it, it must be directly licensed by a following vowel; as in GP failure to be licensed is not an option. In 'sem' this clearly applies to the final /m/. In common with GP I assume that this /m/ is in fact an onset and therefore coordinates in the same way with the previous vowel as does the intervocalic stop in 'gata'; it does not close the preceding syllable, which therefore has a long vowel rather than the short vowel which would be expected in a closed syllable; and it is directly licensed as an onset, and directly governed by, a following empty vowel. The final vowel of 'sem' is genuinely empty, containing no information as to its constituent gesture(s), their constriction location or degree, and consequently it has no internal duration. All we know is that a segment is present which directly governs and licenses the /m/ into an onset and which thus behaves as if it were a vowel. Again as in GP, in certain circumstances these empty vowels must surface, as we will see below in the discussion of syncope, though the manner in which these vowels surface is very different to that of GP. We can represent 'sem' as in (40), where the empty vowel is represented by v°. A final important point is that as v° has no physical content, it cannot indirectly govern the /m/, but the absence of physical content makes indirect government unnecessary. The /m/, however, is indirectly licensed through its relationship with the preceding vowel and this is sufficient to ensure that the /m/ is fully integrated into the prosodic structure.

(40)



Before moving on, let me discuss briefly the structure of diphthongs in AP, as these present a number of problems for which at the moment I have no definite solution. The BNC in Icelandic holds not just for simple vowels but also for diphthongs, so that a stressed open syllable will contain a long diphthong while a stressed closed syllable will contain a short diphthong, resulting in the variation seen in e.g. [fei:t^h] vs [feiht]. In 'gata' [ka:t^{h}a] the single TB gesture for the stressed vowel is the sole content of each of the vowels, head and dependant, and this will derive from a general condition in Icelandic that while a headed vowel can directly license a dependent vowel, the gestural content of each must be identical.[28] In 'feit(t)' however there are two gestures, one for /e/ and another for /i/, and I assume that the condition on the gestural content of head and dependent vowels applies both to simple vowels and to diphthongs. If this is the case we must determine whether we can interpret long simple vowels and long diphthongs in the same way.

We can contrast the short diphthongs of Icelandic with languages which contain short diphthongs in open syllables, e.g. East Perthshire Gaelic (Ó Murchú 1989). For the diphthong in e.g. 'bidhidh' [bei] (future tense, verb 'to be') the two gestures are part of a single segment and the pattern of coordination of the gestures would presumably be comparable to that of the gestures in the labial-velar stops discussed in chapter 3, and subject to the same degree of variation. In Ibibio the velar gesture in /kp/ is weaker than that for /k/ alone. Much the same is true for /gb/ in Igbo, to such an extent that /gb/ and /b/ are very close in duration and [b] is a possible realisation of /gb/. Again, for a prenasal such as /Nt/ possible realisations are e.g. [nt] ~ [^{n}t] ~ [n^t], where all of these are phonologically identical but differ in their physical realisation. These examples highlight the number of factors influencing the physical realisation of all phonological structures. If we take a segment such as /kp/ across a number of different languages, we find that each of them show different degrees of overlap between the gestures, differences in the canonical size of the gestures, their stiffness and target points, the degree of diminution that they show and so on. Nevertheless we can refer to all of these languages as having essentially the same segment despite these differences. In the same way, we regard /s/ and /t/ as both containing glottal opening gestures which are

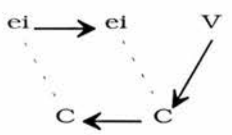
---

[28]This may be a universal rule, though there is no space here to investigate this in detail.

phonologically identical despite the fact that the degree of glottal opening will be in general far greater in the former than in the latter.

While we may speak then of a short diphthong /ei/, the actual physical realisations may differ significantly cross-linguistically, with [e], or [i], or neither dominating. Icelandic short diphthongs, however, always have the structure of long vowels, that is the gestures for the short diphthong /ei/ in Icelandic always constitute both a head vowel and its dependant. In other words, they are in fact long diphthongs (41). At first blush this might suggest that just as the TB gesture for /a:/ is effectively doubled, so the TB gestures for /ei/ are similarly doubled. If this were the case we would expect surface realisations for the long diphthong such as */eiei/, which of course we do not find, although such a realisation would provide us with the correct form for the diphthongs in closed syllables, a coda consonant coordinated with the head vowel both overlapping and hiding the dependant.

(41)



This interpretation is, however, misleading. Whereas the gestures in East Perthshire Gaelic [bei] form a single segment, those of Icelandic form two separate segments, albeit two segments which share a close phonological relationship. We cannot automatically expect coordination of the gestures to be the same in both cases. How then do the gestures combine to form these two segments? Let us first compare the coordination found between adjacent consonants. Within an onset such as /tr/ the two consonants are in a head-dependent relationship forming a single constituent, and this is reflected in the fact that they overlap to a considerable degree. By contrast the consonants in a coda-onset sequence such as /r t/ show much less coordination, reflecting the fact that although the coda /r/ is directly licensed and governed by the following /t/ it is itself also a head. Onset head and dependant thus form one kind of constituent while coda and onset form a different kind, a less tightly bound one.

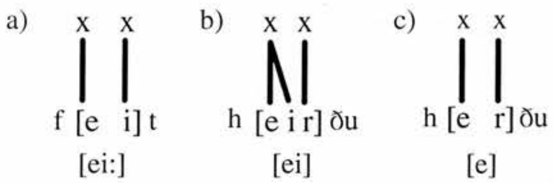
Vowels behave in much the same way. If we had a vowel-vowel sequence, where the vowels were in consecutive syllables separated by hiatus, coordination between the two would be loose, with presumably the offset of the first vowel coordinating with the onset of the second. Where the vowels are in a head-dependent relationship, however, they will show a greater degree of overlap, presumably offset to target. For the long diphthong of 'feit' then we might expect that closer coordination between the two gestures - that for /e/ and that for /i/ - is of this kind. Whereas for /a:/ there is only one head, for diphthongs there are two heads with the accompanying demand that there be two events, the dual headed structure demanding ordering of the gestures. The same principles apply equally whether the gestures form one segment or two, the only difference being in the type of coordination which the gestures show. We can thus expect the two TB gestures for both [ei:] and [ei] in Icelandic to coordinate in the same way as the gestures in the [a:] of 'gata'.

This of course does not automatically imply that the duration of each gesture is identical. As we saw above there is no guarantee that the duration of a gesture when alone will be comparable to the duration of the same gesture when combined with other gestures. Individual languages will undoubtedly vary as to the relative duration of the head and dependent vowels, whether they be simple vowels or diphthongs. This illustrates the natural variation in gestural coordination which we expect to find. Unfortunately, there is no data regarding the realisation of Icelandic diphthongs which would allow us to be more precise regarding the coordination of the component gestures, the two main phonetic sources, Garnes (1974) and Pétursson (1974), not being directly concerned with this matter. The representations [fei:t^h] and [feiht] suggests perhaps that it is the [i] which is lengthened rather than the [e], but this may be misleading.

The phonological relationships involved between vowel segments and following consonants pose a number of problems. Although the diphthongs of

Icelandic usually do surface as such, Pétursson (1974) notes that all of the short diphthongs of Icelandic can undergo a process of monophthongisation as in (42). This is particularly common for /ai/ before /t^h/, though less common for other diphthongs. Within a x-slot or moraic theory of syllable structure, the short diphthongs of Icelandic would be contour segments occupying a single timing slot. Monophthongisation would then involve a failure of the second portion of the diphthong i.e. [i] to associate to a timing slot, whereupon it would delete.

(42)



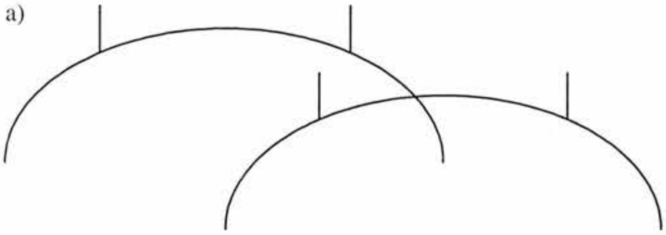
While this would produce the correct result it does little to explain the alternation between contour segments and monophthongs, and there seems no real reason why the deletion should take place. In the closely related language of Faeroese a similar process occurs, though here it is obligatory. Arnasson (1980) reports that there are five vowels which alternate between diphthongs in what he refers to as the long environment, and monophthongs in the short environment. We could assume again that in closed syllables the second portion of a diphthong, whether underlyingly a contour segment or not, is deleted in closed syllables. The analysis of Faeroese would be simpler, however, as there are no short diphthong - short monophthong alternations to account for.

Given the existence of two vowels in long diphthongs, we can expect for [ei:] that the gestures coordinate offset to target, as already noted. Given the presence of a coda in 'feitt' the coda would coordinate with the head vowel, and thus presumably hide in the process the dependent vowel, resulting in a realisation [feht]. While such forms do exist, this does not explain the far commoner realisation [feiht]. Is it possible to explain the presence of such forms within the structures outlined here, or does this reveal a fundamental flaw in our analysis?

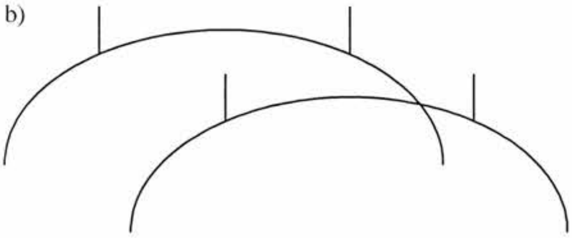
We can represent the coordination of the gestures in a long vowel as in (43) below. While coordination is offset to target, we recall that we refer here to areas in the underlying cycle of the gesture. In (43a) the dependent vowel coordinates a point relatively early in its target with the head vowel's offset. In (44b) on the other hand the dependent vowel coordinates a point relatively late in its target with the same point of the head vowel. This difference between the two will result in a relatively longer vowel in the former, and a relatively shorter one in the latter. Long diphthongs will coordinate in precisely the same way, showing the same degree of variation. Languages will vary cross-linguistically as to the precise kind of coordination they have and what type of variation is possible within the language. Matters grow more complex once a coda is present, as while a coda will coordinate its target with the offset of the head vowel, the same type of variation as to how this coordination is realised will be expected. The combination of these factors will determine the surface form of diphthongs in open and closed syllables.

(43)

a)



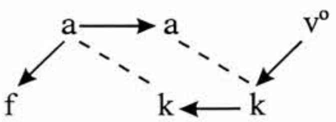
b)



The precise pattern of coordination in Icelandic would need to be determined instrumentally, but we can now point a way to a solution of the problem of the short diphthongs. (43b) makes the point most clearly, though I do not suggest that this is the actual pattern found in Icelandic. If the coda /r/ in 'heirðu' were to coordinate with the offset of the head vowel in (43b) this would mean that the initial period of the head vowel's target would be left uncovered, but in addition much of the target of the dependent vowel would also be left uncovered, resulting in a surface form [heirðu]. If the coordination of the vowels were now to alter so that the target portions showed considerably less overlap as in (43a), the same coordination with the coda /r/ would now result in a surface form [herðu]. We thus achieve the desired goal of describing the alternation between forms with diphthongs and those with monophthongs without recourse to any rules of deletion or of changes in segmental structure, employing only a change in the coordination of the vowels within an expected range. In Faroese there is no such change of coordination, the vowels in the alternating forms showing coordination more comparable to that in (43a), the presence of the coda then simply hiding the dependent vowel.

As might be expected, the structure of geminate consonants is directly comparable to that of long vowels. In 'telpa' in (39) the onset directly and indirectly licenses and governs the preceding coda, and the two positions then coordinate as described. In a word such as 'flagg' [flak:] (flag), however, the coda and the onset are identical, in other words there is a geminate. We can describe such structures straightforwardly as in (44), where the onset directly licenses a coda position as it does in 'telpa'. The difference between the two lies in the fact that in (44) the onset licenses itself to occupy both positions, just as the stressed vowel licenses itself to fill the dependent vowel created by the BNC in the same word. The geminate consonant in [flak:] is an unaspirated stop, so that it contains only one headed gesture. In this sense it is directly comparable with the nucleus of [ka:ta] and both can be realised in precisely the same way, that is by effectively realising the entire segment twice. How though does this analysis affect the realisation of aspirated geminates?

(44)



The structures of GP demanded that there be a general constraint that the friction element h° always spread, and that was this spreading which gave rise to both post- and preaspiration. The structures of AP mean that no such rule is necessary. Given the dual headedness of aspirated stops, offsetting of the gestures is obligatory, automatically resulting in postaspiration in e.g. 'gata' [ka:tʰa]. Similarly in 'titra' [ti:tʰra] postaspiration is present and the glottal opening gesture will automatically overlap the following sonorant, devoicing it in the process without any process of spreading. The absence of such a spreading rule avoids the problems noted above for GP in forming light diphthongs consisting of the friction element h° and a following vowel. At the same time, we must recall that there is in fact nothing in the structures developed in chapter 3 which indicated the actual ordering of the component gestures in an aspirated stop, the same being true of other complex segments such as /gb/ and /ts/ where the ordering seems to be fixed universally, and of oral-nasal segments where prenasals are very much commoner than postnasals.

For aspirated stops, only a very few languages, of which Icelandic is one, prevent us from making the universal claim that aspirated stops always order the oral closure before the period of unaccompanied glottal opening. Thus, although in the structures argued for here the oral and glottal gestures in aspirated stops are both heads and therefore theoretically of equal status, it does appear as if in reality the oral gesture controls the glottal gesture in some way so that the glottal gesture is overwhelmingly ordered after the oral closure. However, the reasons behind this remain obscure, as we saw in the discussion of Kingston (1990) in chapter 3, and the existence of e.g. both prenasals and postnasals means that we cannot simply ban one ordering or the other, leaving us only able to say that for any particular language there is never a phonological contrast between the two possible orderings.

Given that gemination of the unaspirated stops results in simple lengthening of the stop, we would initially expect the same to be true of the aspirated stops, giving us forms such as *[tappʰi]. This is what happens in closely related languages such as Swedish and Danish. Instead, of course, the surface form is [tahpi], with the oral closure remaining occupying the onset position but with the coda position occupied solely by the glottal opening gesture. This is consistent with the demand that the two headed gestures be offset, i.e. the gestures are offset whether the stop is a singleton or a geminate, but the ordering in geminates is the reverse of that found in singletons. Geminate nonaspirates and geminate aspirates thus appear to differ from each other in the same way as do simple long vowels and long diphthongs in terms of their internal structures. Accordingly they behave in similar ways. Mimicking the pattern seen above for long diphthongs, one of the headed gestures occupies the coda position, the other the onset. In Icelandic it is the glottal gesture which occupies the coda, and at the same time continues to overlap the oral gesture which itself occupies the onset position. The resulting geminate is realised as [hk], with pre- rather than postaspiration.

The glottal gesture need no longer extend beyond the oral gesture to produce a second event as its syllabification into the coda automatically fulfils the requirement for two events. Browman & Goldstein (1989), following Thráinsson (1978), suggest that preaspiration is achieved by deleting the supralaryngeal gestures of the first of two homorganic stops, e.g. kk ---> hk, without any consideration of the effect on syllable structure. Instead we can see that aspirated stops behave in basically the same way as other geminates and in a way comparable to prenasal geminates. It is their dual headed status that allows aspirated stops to be realised as preaspirates when geminated, with no need for any deletion of gestures or arbitrary processes of lengthening.

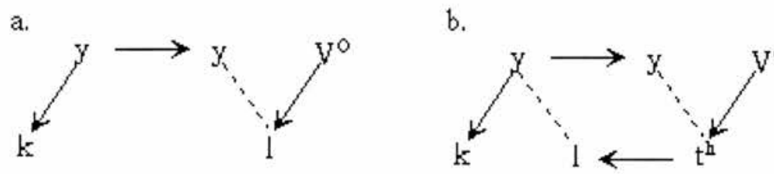
This is comparable with the analysis of Sinhalese geminate prenasals in chapter 3, where gemination of /^{n}d/ resulted in /nd/, the major difference being that in Sinhalese the ordering of the two gestures is the same in both forms. There is apparently nothing in the phonetics of Icelandic which provides any clue as to precisely why preaspiration is chosen over the overwhelming crosslinguistic preference for postaspirates, though we shall see a preference for the same ordering in other environments below. This also suggests another conceivable realisation of a geminate aspirate, namely /kh/ i.e. a heavily aspirated stop, and while I know of no languages with such a segment, the segmental and syllabic structures developed here suggest that such forms are possible. It may be that such forms do exist but have been given another interpretation, e.g. a fortis stop. This clearly requires further investigation.

We might describe preaspiration as a process whereby the headed glottal gesture of aspirated stops occupies the coda if one is present, but forms such as [tɛl̥pa] in Southern dialects suggest an alternative description. In [tɛl̥pa] the coda sonorant is devoiced, suggesting a process of spreading of the glottal gesture to the coda consonant, but rather than viewing this as a process of spreading to a segment I suggest that it is best viewed as spreading to the coda position. In other words, if an aspirated stop in the onset indirectly licenses a coda - which is the same as saying if it coordinates with a coda - then the glottal gesture of the stop will itself occupy the coda position, i.e. it will coordinate with the preceding nuclear head. The result is that preaspiration will result whether or not the stop is a geminate, so that spreading of the open glottis therefore takes place regardless of the presence of other segments in the coda. Again, we can compare this with the gemination in Havana Spanish discussed in chapter three, where the informal speech realisation of /kurba/ as [kubba] is achieved by gemination of the stop but without deletion of the /r/ which is simply covered up and never heard. A still simpler comparison is with the casual speech devoicing of English voiced obstruents before voiceless obstruents e.g. cab̥ fare.

For Southern dialects this is a sufficient description, as all codas preceding aspirated stops are voiceless. Northern dialects, on the other hand, do not show this devoicing, 'telpa' surfacing as [tɛlp^{h}a] with a postaspirated stop as opposed to the [tɛl̥pa] of Southern dialects. We can assume that spreading in these dialects is simply blocked when the coda is already occupied by other gestures. In fact, these dialects do show some devoicing of codas, but this is of the same type found in English, and is not caused by preaspiration so that these consonants are voiceless before any voiceless consonant and not just aspirated stops. Southern and Northern dialects differ, then, in that the former has extended and phonologised the tendency for the glottal gesture of aspirated stops to spread to the coda.

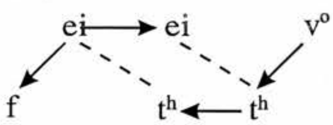
We can apply the same principles to forms which show variation between long and short vowels, as in the feminine and neuter forms of adjectives. For example, 'gul' [kY:l] (yellow) has a long vowel due to the BNC. The word-final /l/ has no following consonant which could govern it and license it as a coda, so it must instead occupy an onset position (45a). We can think of this as the default syllabification, where since syllabification of /l/ into a coda is not possible it must of necessity be in an onset and must be followed by an empty nucleus. The addition of the neuter suffix results in 'gult' [kYlt^h] ~ [kYl̥t], with a short vowel (45b). As the neuter suffix is now able to govern the /l/ and license it straightforwardly into a coda position in precisely the same way as in 'telpa', there is no longer any need to assume the presence of an empty nucleus licensing /l/. Instead the suffix itself now occupies the onset position. This is not, of course, the only possible interpretation, however, and other languages may instead have unchanging underlying forms.

(45)



The same principles apply again to forms such as 'feitt', with the exception that two identical consonants seem in all environments to be interpreted as true geminates (46). This mirrors Gibb's account of a violation of the OCP, except that here a sequence of two identical consonants, one of which occupies an onset and licenses the other into a coda, is interpreted as a single segment. The addition of the neuter suffix to 'feit' thus creates a geminate which again has the same structure as that in 'kappi'. By occupying both the coda and the onset the aspirated stop is realised as a preaspirate. By assuming that all such sequences are interpreted as geminates we avoid the problems of the GP analysis.

(46)

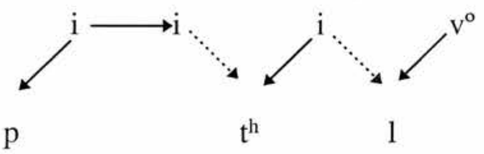


Finally, the neuter suffix triggers spirantisation in both /k^h/ and /p^h/, as we have seen in 'rakt' [raxt]. If mentioned in relation to preaspiration at all this rule is generally stated as one which has the effect of leniting /k^h, p^h/ to the corresponding spirants before /t^h, k^h, s/ (with the natural exception that /k^h/ followed by another /k^h/ results in a geminate). We can see this in 'skips' [scɪfs], genitive of 'skip' (ship). However, spirantisation is dependent not just upon the presence of the putative triggering consonant but also upon syllabification. Compare the alternative form [scɪ:ps] where the stop remains unlenited and the vowel is not shortened, despite the presence of the supposed triggering environment in both cases. This would suggest that the genitive suffix -s may alternatively not syllabify preceding consonants into the coda, so that the labial stop in [scɪ:ps] remains in an onset. For the stop to be spirantised, then, it is syllabification in the coda, rather than the identity of the following consonant, which is the trigger, confirming that the syllabification of the final consonant of the stem of both 'heilt' and 'rakt' (/l/ and /k^h/ respectively) into the coda is the correct analysis. I shall return to this below.

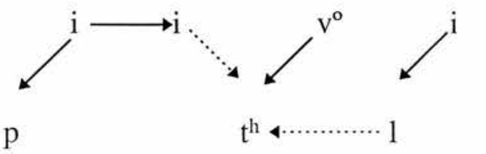
I turn now to preaspiration in syncope environments. Given the principles of syllabification above, 'Bitil' has a straightforward form (47a). The final /l/ is directly licensed by an empty vowel, but this vowel is unable to indirectly license it. As all consonants must be indirectly licensed by a vowel, /l/ must be indirectly licensed by the preceding nucleus, as we saw earlier for 'sem'. The resulting form is [pi:t^hil] with syllabication following the expected pattern. The addition of the dative singular suffix provides a full vowel to which the word-final /l/ can associate, and we would initially expect *[bi:tili] as the surface form. However, the additional vowel seems to act as the trigger for syncope, causing the preceding vowel to be inaudible. Browman & Goldstein (1986) provide a simple gestural account of syncope in English. They note that 'beret' may in some environments be realised as [bərei], in others as [brei] (Am. Eng.) where the initial vowel is elided. This can be simply represented in gestural terms as a process of overlap where the initial /b/ and the following /r/ gradually slide over the intervening unstressed vowel, eventually overlapping each other and in the process overlapping the vowel, which is then inaudible though still present (and perhaps diminished to a greater or lesser degree). A further example is 'secretary' with the possible realisations of [sɛkrətɛri], [sɛkrətəri] and [sɛkrtri]. Here the /ɛ/ in the third syllable becomes gradually less stressed, more reduced and more overlapped by the flanking consonants until it eventually is completely hidden. This process is not of course necessarily categorical and we might then expect to find varying degrees of both consonant to consonant overlap and overlap of the vowel. Stressed vowels in general are not subject to syncope, and there is clearly a connection between absence of stress and the amount to which a vowel can be overlapped by flanking consonants. However, the study of the metrical structure of Icelandic lies beyond this thesis.

(47)

a)



b)



Non-compound words in Icelandic have the main stress on the initial syllable, with secondary stress on the third, fifth and so on, all other syllables being unstressed. In 'Bitil', then, the second syllable is unstressed. Once a third syllable is added, as in the addition of the dative singular suffix /i/, the second unstressed syllable of 'Bitili' would now fall between two stressed syllables.[29] Historically, unstressed vowels in this environment were, I suggest, subject to the same processes of gradually increasing overlap by flanking consonants as in English, but that what was originally a casual speech phenomenon is synchronically categorical. If increasing overlap of the vowel by the consonants had a similar effect as it has in English 'secretary' then we could expect the unstressed /i/ to weaken, perhaps to schwa or weaker still until it was completely hidden and hence inaudible. I follow Gibb in claiming that the unstressed /i/ in (underlying) 'Bitili' is realised as an empty vowel, but there seems no need for any process dependent upon interconstituent government, or the lack of it, between adjacent nuclear positions. Rather, it seems more straightforward to avoid the problems of Gibb's analysis and simply claim that unstressed vowels which directly license stops are realised as the empty vowel when preceded by a stressed vowel and followed by a stressed vowel which directly licenses a sonorant. This mirrors the historical development, and while it historically is clearly dependent upon the metrical structure of the language, synchronically it is an idiosyncratic feature of the language and not one which should be accounted for by introducing new forms of interconstituent government. The vowel/empty vowel alternation is then simply lexicalised for those words in which syncope occurs.

This alternation between full vowel and empty vowel occurs everywhere in which its environment is met, resulting in the structure in (47b) for 'Bitli'. The final /l/ is now indirectly governed by the following vowel, and normally we would expect an intervocalic onset head to also coordinate with the preceding vowel. This preceding vowel, though, has no physical content and as the /l/ is already indirectly governed by a vowel there is no requirement for it also to be indirectly governed by the preceding vowel if it is empty. However, the /l/ must still form an indirect governing relationship with the previous syllable, otherwise it will fail to be integrated into the prosodic structure of the word. We must now revise the requirement that onset heads necessarily coordinate (i.e. form an indirect government relationship) with a preceding nucleus and claim instead that an onset head will

[29] As 'Bitil' is a modern loan this process of syncope never actually occurred in this word, though it

coordinate with a preceding nucleus if that nucleus has segmental content, otherwise it will coordinate with a preceding onset, provided such an onset is present. For (47b) this means that there must be an onset with which the /l/ can coordinate, and also that the /l/ is still indirectly governed by a vowel. Both of these conditions are met.

Syncope thus involves coordinating two consecutive consonants, creating an indirect government relationship between the two, and it is this government which triggers preaspiration. We recall that 'tl' is an illicit consonant cluster in Icelandic onsets so there is no reason to expect such a cluster to be licit after syncope, despite Gibb's claims to the contrary. We also recall that interconstituent government in GP seems always to be from right to left, violated only, apparently, by government between Icelandic onsets. I assume that interconstituent government in Icelandic, at least in words such as 'Bitli', does indeed go from right to left. The existence of an indirect government relationship between both /t^h/ and /l/ does not alter the fact that /t^h/ and /l/ are each the head of an onset. If indirect government between separate constituents goes from right to left, the stop as well as being a head is now also indirectly governed by a consonant to its right, an onset. What is the interpretation then in Icelandic of such a consonant, one which is both head of an onset and also indirectly governed by an onset? Elsewhere in the phonology this quite clearly defines a geminate. For example, the intervocalic stop in 'kappi' is simultaneously both head of an onset and governed by one. The result then in 'Bitli' is that /t^h/ behaves here just as we might expect and is realised as a preaspirated stop, its glottal gesture coordinating with the head of the preceding nucleus.

No phonological process of syncope as such is needed for forms such as 'epli' where syncope is fully lexicalised - i.e. 'epli' has the structure /ɛpv°li/ - and no alternative forms exist which contain a phonetically audible vowel between the stop and the liquid. Forms such as 'vopn' are similarly anomalous. Recall Gibb's analysis of such forms (given here again as (48a)). The final empty category is unable to

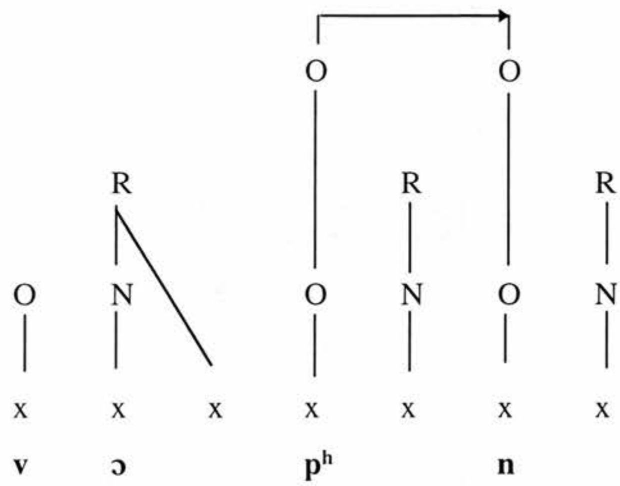
---

did in others.

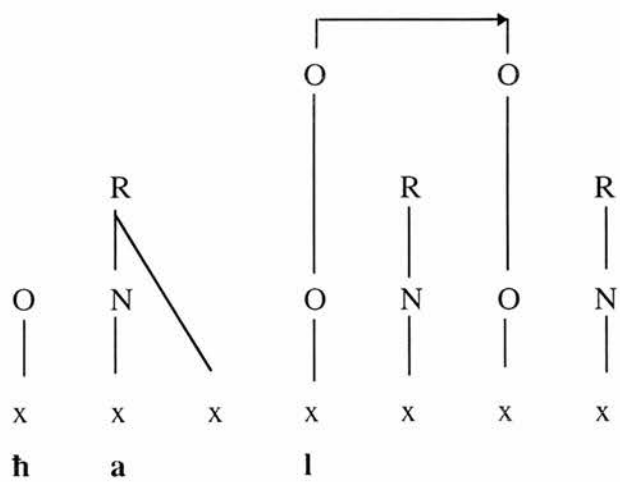
properly govern and thus license the preceding nucleus which we then expect to be phonetically realised. Gibb, however, assumes that the onset to onset government between the stop and the nasal forms a governing domain, allowing the vowel to remain phonetically inaudible. This is based on the analysis of Kaye (1990). Kaye discusses the case of the active participle [ħa:ll] of the verb 'to open' in Moroccan Arabic (48b). Again the final nucleus is unable to properly govern the preceding nucleus which we expect to be phonetically realised. Kaye however suggests that the doubly linked /l/ constitutes a governing domain, and Gibb adapts this to suggest that any onset to onset government acts as a governing domain.

(48)

a)



b)



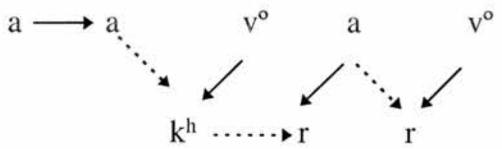
If onset to onset government can occur regardless of whether the intervening nucleus is properly governed then there seems no need to invoke proper government of nuclei as part of the process of syncope. I suggest that forms such as 'vopn' are anomalous in that they allow indirect government between /pʰ/ and /n/ regardless of whether or not the /n/ is followed by a full vowel. The relationship between the stop and the nasal is lexicalised, and thus licensed independently of the presence or absence of a following nucleus. On the other hand, forms such as 'Bitli' are not so licensed and do depend on the presence of a following full vowel. Contrary to Gibb's analysis I suggest that it is 'vopn' which is anomalous within the system, reflecting a fossilised form.[30]

The anomalous status of 'vopn' is reflected in the variations in the surface realisations of a number of words. As already noted, syncope in 'akrar' does not result in preaspiration, and we can assume it has the structure in (49). For 'Bitli' we know that 'tl' is not a possible onset cluster, leading us to assume the presence of an intervening empty nucleus. For 'akrar' the stem form is 'akr', the surface form [a:kYr] arising from a process of epenthesis (Oresnik 1972). Such epenthetic vowels are assumed by GP to be underlying, and I suggest the same is true of AP where the /kʰ/ and /r/ both occupy onset positions and are licensed by empty vowels. In 'akrar' the addition of a full vowel again results in the presence of an indirect governing relationship between the liquid and the preceding stop, causing the intervening nucleus to remain inaudible. However, whereas 'tl' in 'Bitli' is blocked from being interpreted as a licit cluster i.e. /tʰ/ cannot govern /l/, 'kr' in 'akrar' is explicitly specified as being a licit cluster; that is when a stop and /r/ are physically adjacent they are interpreted as a genuine cluster, regardless of any intervening vowels. Thus the stressed nucleus surfaces as long and no preaspiration occurs. In 'akur' the lack of a following full vowel results in the failure to license syncope and the preceding nucleus is realised with a full vowel. Why should this be so?

---

[30]We can compare 'vopn' directly with 'vopna' (to arm) where a final nucleus is available to license the syncope. If we assume that such forms are related in some way then they are directly comparable with forms such as 'pukr' discussed below. Of course, there remains the question of precisely how the relationship between /pʰ/ and /n/ is licensed. This presumably takes place at the level of the word, but this needs to be established.

(49)



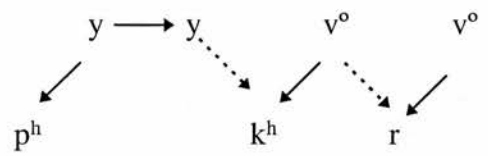
In the forms discussed so far all segments within the word are coordinated with other segments. The need for such physical coordination gives rise to the coordination between the onset head and the immediately preceding vowel, this coordination resulting from the indirect government by the vowel of the onset which serves to fully integrate the onset consonant into the prosodic hierarchy. Such coordination of a syllable with a preceding one is obligatory. However, forms such as 'Bitli', 'epli' and 'vopn' make it clear that coordination does not necessarily occur between consonant and vowel but may also occur between two consonants. If a full vowel precedes with which the consonant can coordinate, then coordination with this vowel is obligatory. For words derived from the stem 'akr-', onset to onset coordination rather than onset to nucleus occurs providing a full vowel is present which can indirectly govern and thus license the final /r/. However, for 'akur' no such vowel is present so the final /r/ must coordinate with the preceding vowel in order to be indirectly licensed. This coordination will take the usual form, offset of the vowel to target of the consonant, but in order for there to be an offset the vowel must of course contain information as to its gestural content. The empty vowel is therefore required to contain a gesture, the details of which may differ from language to language, and since this gesture is not hidden it is audible. Like GP there is no insertion of a vowel, rather the vowel is always present underlyingly, though it does not always have any internal content. Unlike GP the realisation or not of the empty vowel does not depend in any way upon the identity of the vowel in the following nucleus, other than indirectly in that vowel's indirect licensing of a preceding onset, nor on any nucleus to nucleus government.

Realisations such as [a:kYr] reflect the fact that clusters such as 'kr', 'tr', 'kj' are not found word-finally. However, there exists a small class of derivative nouns formed from verbs (50a-c) which break this generalisation (Arnasson 1980). Initially we might wish to analyse these as containing branching onsets followed by a null vowel, otherwise identical in form to 'akrar'. Some speakers though, Arnasson amongst them, have a tendency to insert an epenthetic vowel into these clusters as in (50d-f). The simplest analysis would be to assume that the epenthetic vowel is again always underlyingly present, so that the structures for e.g. (50a) and (50d) are identical (51), with all of the consonants licensed as onset heads. The difference between the forms with and without epenthesis lies in whether or not the relationship between the stops and /r/ is licensed despite the absence of a following full vowel. The unusual and rare nature of these forms, not all of which, as Arnasson points out, are accepted as licit words by all speakers, can allow them to be treated in the same way as 'vopn' where the syncope is lexicalised and thus exempted from normal licensing considerations. That is, for these words the requirement that consonants must be indirectly licensed by a vowel is relaxed. We can compare this with English 'keep' / 'kept', where Kaye (1990) argues that the change in vowel quality and quantity is no longer due to a phonological rule but is lexicalised. Alternatively, speakers may reanalyse these forms and force them to comply with normal phonological rules where the absence of a full vowel prevents the licensing of government between the stops and /r/ and results in the apparent insertion of an epenthetic vowel. In either case the stressed vowel remains long because the stop is never indirectly governed and thus never able to occupy the coda.

(50)

a.	pukra	'to be secretive'	pukr	[p^{h}Y:k^{h}r]	'the act of being secretive'
			or	or	
			pukur	[p^{h}ʏ:k^{h}ʏr]	
b.	sötra	'to sip'	sötr	[sœ:t^{h}r] or	'the act of sipping'
			or	[sœ:t^{h}ʏr]	
			sötur		
c.	kjökr	'to wail'	kjökr or	[c^{h}œ:k^{h}r]	'the act of wailing'
			kjökur	or	
				[c^{h}œ:k^{h}ʏr]	

(51)



Arnasson points out that similar variation in syllable structure is also found in compound words. For example, 'torfæra' [tʰɔr̥fai:ra] has the short /ɔ/ of a non-compound word, but the voiceless [f] of a compound. The existence of such variation in syllable structure is crucial to the correct interpretation of the forms in (52), where 'litka' is an alternative form of 'lita' (colour), which contain a sequence of two non-homorganic stops. The forms in (52) are taken from Oresnik & Pétursson (1977), who suggest that vowel length is unaffected by syllabification, citing the first two realisations in support of this. They are unable however to account for alternative realisations and simply discount them.

(52)

a. [lɪ:t^{h}k^{h}a]

b. [lɪ:tka]

c. [lɪh(t)ka]

d. [lɪθka]

(52a,b) are presumably Northern and Southern forms respectively, and contrary to expectations we see that the stressed vowel is not shortened by the addition of a consonant-initial suffix. Realisations such as these suggest that 'litka' is here treated as if it were a compound, where the suffix is in a different cycle and thus unable to license the preceding stop into the coda.[31] We can compare these with a clear compound such as 'litlaus' [lɪ:tl̥œys] (colourless) where the stressed vowel is similarly long. (52c) on the other hand suggests a parsing where 'ka' is treated as a normal inflection, creating a syncope environment and governing the preceding onset. Preaspiration is the expected result.[32] The final form is the most interesting of all. We recall the rule of spirantisation where /p^h, k^h/ are realised as /f, x/ apparently before another non-homorganic voiceless stop or /s/. Although the realisation of 'litka' as [lɪθka] is uncommon it does occur, yet the spirantisation is generally ignored or at best considered an unimportant anomaly. The fact that the stressed vowel is also shortened again suggests that this form is interpreted as involving simple licensing of /t^h/ directly into the coda, spirantisation again being dependent not on the identity of the following segment but on syllabic structure in general and the syllabification of the stop in particular.[33]

The structures put forward here clearly share much in common with the structures of both GP and DP, but differ from them in significant ways. In particular, the interpretation of licensing and government is distinct from that of either theory. Direct government encodes constituent status within the theory, so that effectively three primary constituents are recognised, namely the nucleus, the onset and the coda. Each of these are characterised as minimally consisting of a head which may directly license a dependent, non-head. In other words, each of the three constituents may branch. Indirect licensing serves a quite different function, acting to integrate the various segments within the word into the prosodic hierarchy. Every segment within the word which has physical content must coordinate with at least one other

---

[31] This calls to mind the behaviour of the genitive /s/ suffix discussed above.

[32] The presence of preaspiration is compound words is commoner in some dialects than others. Pétursson (1974) notes that such realisations are commoner than elsewhere in the dialect of Austfirðir, where e.g. 'laklega' (slowly) is realised as [lahkleɣa] rather than [la:k^hleɣa].

[33] Icelandic spirantisation is discussed in chapter 5.

segment within the word, provided that the word consists of more than a single segment, and this coordination is achieved through indirect licensing. Not all segments have physical content, however, and this plays a crucial role in the analysis of epenthesis and syncope of Icelandic.

While much of the structure outlined so far is universally applicable, it is obviously undesirable to base a theory of syllable structure on the basis of the study of a single language. Indeed, there is evidence that not all languages share all of the same basic structures, and it is to this that I now turn.

4.4
Italian Syllable Structure

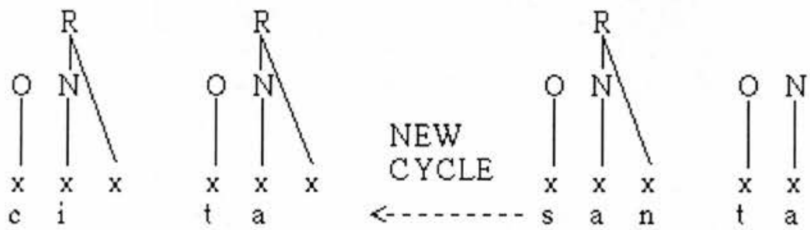
Italian is cited by Gibb as another language which shares with Icelandic the BRC and which shows similar phonological consequences. Stressed syllables branch with the rhymal point created by the BRC again apparently filled by either a consonantal or a vocalic segment, so that such syllables contain either a long vowel, or short vowel plus consonant. For example, the BRC converts 'fato' (fate) into [fa:to] and 'fatto' (fact) into [fatto]. Again Gibb assumes that in 'fato' it is the Cyclic Spreading Constraint which acts to prevent the following stop from spreading into the available rhymal slot.

As further evidence that the CSC is active in Italian as well as Icelandic Gibb discusses the process of Radoppiamento Sintattico (RS) whereby the initial consonant of a word is syllabified leftwards into the coda of a preceding word when the preceding word ends in a stressed vowel (53), leading to gemination of the consonant (except in s + C clusters, where the /s/ is simply syllabified into the coda without gemination). In 'cittá santa' Gibb assumes that the BRC applies to the final stressed vowel to create a rhymal slot just as it does in 'fato'. The slot is then filled by the initial /s/ of the following word, which being in a separate cycle is not bound by the CSC, much as the neuter suffix in Icelandic is free to spread. When no consonant is available the nucleus is claimed to spread, again as in Icelandic (54).

(53)

a.	paltó pulito	[paltoppulito]	'clean coat'
	é carino	[ekkarino]	'it is pretty'
b.	cittá triste	[tʃittattrist]	'sad city'
	caffé freddo	[kafeffredo]	'cold coffee'
c.	cittá santa	[tʃittassanta]	'holy city'
	caffé serale	[kaffesserale]	'evening coffee'
d.	cittá straniera	[tʃittastraniera]	'foreign city'
	caffé spesso	[kaffespesso]	'thick coffee'

(54)



There are a number of problems with this analysis. First of all, the connection between vowel length and RS is tenuous. In addition to occurring after polysyllables with final stress, RS is also triggered by stress bearing monosyllables such as 'da' (gives), unstressable monosyllables such as 'e' (and) and penultimate stressed polysyllables such as 'come' (like) (Loporcaro 1988), all such words ending in short vowels. Stress and its concomitant lengthening therefore are not prerequisites for RS. More serious is the fact that in final stressed words such as 'cittá' stress does not, in fact, ever result in a long vowel, these words always surfacing with a short vowel, so that the default spreading proposed by Gibb of the final stressed vowel in to the metrically created rhymal slot is erroneous. We could account for this fact by blocking spreading of the vowel when the syllable is word-final, but this would be ad hoc and arbitrary.

While it is true that stress in Italian does behave somewhat like that in Icelandic, there are clear differences. An initial generalisation which we can make is that, within the word at least, in order for a stressed vowel to be long it must be followed by a consonant. We might then postulate that some relationship is created between a stressed vowel and a following consonant which causes the vowel to surface as long, so that the absence of the consonant results in the absence of the relationship and therefore the failure of the vowel to be long. If this is the case, what might such a relationship be and what implications would it have for syllable structure?

Farnetani & Kori (1990) show that stressed vowels in Italian tend to be longest in penultimate position so that, as we might expect, gradient factors also influence the overall duration of stressed vowels. Discounting this, we have assumed so far that in a gestural approach a vocalic gesture will coordinate its offset with the target of a following consonantal gesture. The tail of the vocalic gesture is inevitably hidden somewhat by the consonant, and by 'delaying' the onset of the consonant we can reveal more of the vowel, resulting in the vowels duration increasing. Conversely, by beginning the onset of the consonant earlier we hide more of the vowel and decrease its duration. As noted earlier, these processes must entail viewing the target and offset of gestures not simply as single points in the gestural cycle but as larger overlapping areas. When we speak of delaying a post-vocalic consonantal gesture we imply coordinating its onset with a slightly later point of the vowel's offset area.

In an analysis of vowel length in Italian nonsense words Dunn (1990) suggests that stressed vowels in closed syllables undergo closed syllable vowel shortening, achieved by shifting the consonant to an earlier point in the production of the vowel. The relationship between the stressed vowels of 'tapa' and 'tappa' must then be one in which each has underlyingly the same long stressed vowel, with that of 'tappa' shortened from being overlapped to a greater extent by the following geminate. What this fails to answer is how it is that the vowels are long in the first place, and how they differ from the short vowels in the final syllables.

Crucial to Dunn's analysis is the fact that there is no process of lengthening, only one of shortening. Comparing the stressed vowels in 'tapa' and 'tappa' Dunn found that the amplitude of vocalic opening peaks earlier in closed syllables, amplitude of lip aperture is reduced, and the vowel's duration as a whole is much shortened. However, if there is indeed some relationship between the stressed vowels and their following consonants which results in long vowels, we would expect there to be an active lengthening process. Such a process would be equally compatible with the phonetic phenomena noted by Dunn, only viewed from the other side as it were, where such lengthening would increase the vowel's duration, increase the amplitude of lip aperture, and peak amplitude of vocalic opening would occur later in open stressed syllables.

Recalling the work of Beckman et al (1992), they noted that the lengthening of vowels in American English stressed syllables was accomplished by decreasing the amount of overlap between the stressed vowel and a following consonant, so that less of the vowel was hidden, thus automatically increasing its length. This carries with it the implicit notion that if all of the vowel were to be uncovered, as would be the case in a word containing only a single stressed CV syllable, length would be maximal. While vowels in such syllables may be in general somewhat longer than those in closed syllables, the difference would be unlikely to be enough to suggest a categorical difference in vowel length. When a vowel is stressed, it appears from Beckman et al's study that the achievement of a following consonant's target is delayed with respect to that of the vowel's offset. We saw above the variation possible in realising offset to target coordination for vowels, and vowel to following consonant variation was noted as also showing the same degree of variation. Such variation alone could account for much or all of the difference in duration between stressed and unstressed vowels in American English.

If we were to continually realise a consonant later with a preceding vowel, coordination between the two would eventually cease. The minimum coordination the two could show would be to coordinate the offset of the vowel with the onset of the following consonant. The resulting change in length would again be little greater than for a stressed vowel with no following consonant. What would considerably lengthen the vowel would be if such a gradient shift were not simply to replace one type of coordination with another but were instead to add an additional type of coordination. How can this be achieved?

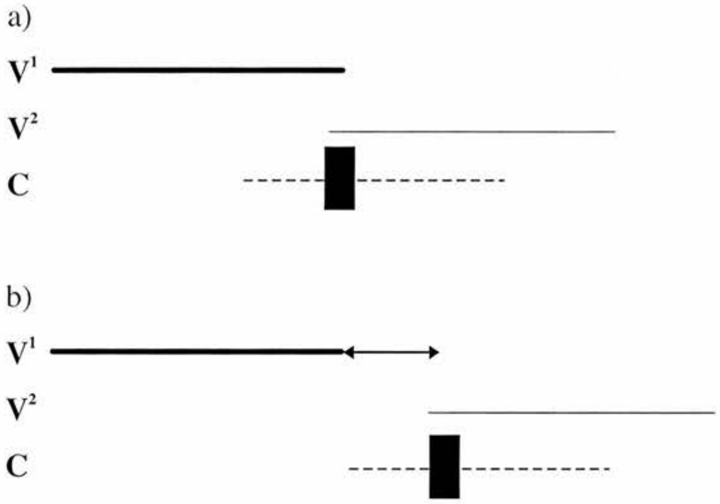
Smith (1995) provides a solution to this question. She proposes a model of syllable structure for Italian which differs from that proposed here for Icelandic in that rather than the vowels of separate syllables coordinating with each other only indirectly through their coordination with an intervening consonant, she suggests that they instead coordinate directly as in (55) where both vowels are heads (reinterpreted in terms of the relationships proposed here). This has a number of implications. Firstly, it means that, unlike Icelandic, onset heads in Italian are not required to enter an indirect government relationship with the preceding syllable, and secondly it means that vowels on the other hand must enter an indirect government relationship with a following vowel if one is present. Simplifying, whereas Icelandic requires coordination of syllables to run from left to right, Italian requires it to run from right to left.

(55)

$$\begin{tikzcd} V^1 \arrow[r, dotted] & V^2 \arrow[dl] \\ C & \end{tikzcd}$$

Vowel to vowel coordination, where V^2 is not directly licensed by V^1, coordinates the offset of V^1 with the onset of V^2. As V^2 also coordinates its onset with the target of the consonant in (55), the consonant's target is thus indirectly coordinated with the offset of V^1. If the effect of stressing V^1 were to directly coordinate it with the following consonant as well as the following vowel, no significant change would occur were it to coordinate with the target of the consonant. Were a stressed vowel instead to coordinate with the target of the following consonant, the effect would be to lengthen that vowel in the way noted by Beckman et al. This lengthening is illustrated in (56). In (56a) the two vowels minimally overlap, and both overlap the target of the intervocalic consonant, as represented by the solid box. In (56b) coordination of V^1 with the onset of the consonant causes V^1 to lengthen, the additional length shown by the double arrowed line. While this extra length may appear to be rather small, Beckman et al. show that for English this lengthening be upwards of 100msec, more than enough to account for the difference in length between Italian long and short vowels.[34]

(56)



This structure for the stressed vowels in Italian has two immediate effects. It automatically uncovers the period of V^1 which would otherwise be hidden by the consonant, and in addition delays the offset of the vowel, prolonging the vowel's target and thereby effectively stretching it beyond its normal duration. We also maintain vowel to vowel overlap, whereas if we simply shifted coordination of the vowel's offset to e.g. the consonant's onset we could not guarantee that such overlap

[34] It is generally considered that phonemicisation of such a phonetic process was responsible for the shift to stress-induced vowel lengthening during the change from Vulgar Latin to Italian (Elcock 1975).

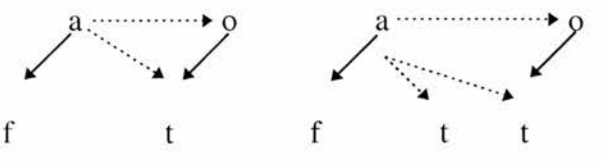
would always be maintained. The effects noted by Dunn simply reflect less overlap plus a greater amount of time in which the vocalic gesture can reach its goal.

The effect of any BNC is thus not to create an additional dependent vowel but rather to create an additional vowel to consonant relationship. In final stressed words such as 'cittá' the absence of a following consonant of course means that no vowel-consonant relationship is possible. This does not mean that the final vowel in 'cittá' is necessarily of the same general duration as the preceding unstressed vowel or the unstressed /a/ in 'alta', as it is not overlapped by a following consonant and other factors may be at work to lengthen stressed vowels in general, as noted by Farnetani & Kori (1990).

As in Icelandic, closed syllables, e.g. the initial syllable of 'alto', contain underlyingly long vowels. The onset head /t/ directly licenses the /l/ into the coda as in Icelandic, but unlike Icelandic there are not two vowels with which the consonants can coordinate. The stressed vowel's offset is coordinated with the onset of the /t/ and the onset of the /o/, but now also with the target of the /l/ as in Icelandic. While this may seem to be a complex maze of coordination, we should recall that it is no more than we would require anyway in any phonetic interpretation. We can represent syllable structure in Italian as in (57).

(57)

a) b)



The stressed vowel of 'fato' in (57a) coordinates its offset with both the following consonantal gesture and the following vowel gesture, indirectly governing both. The stressed vowel of 'fatto' in (57b) shows the same coordination but with additional coordination with a coda consonant. If Beckman et al (1992) are correct in their description of vowel lengthening in English stressed syllables as being caused by later phasing of the vowel with a following consonantal gesture, then vowel length in Italian has made categorical something which was possibly once only gradient. Italian and Icelandic, while sharing many features, differ crucially in their syllable structures in a number of ways. These differences provide some answer to such criticisms as those of Clements (1992) that more gestural relationships were possible than were actually attested. Icelandic and Italian, however, while making use of the different types of intergestural coordination possible nevertheless constrain these in terms of the phonological relationships which hold between segments. Given that other types of coordinative relationships are possible between gestures, it may very well be that there are still more types of syllable structure possible, which may differ more or less from those of Icelandic or Italian. In addition, it is also feasible that a single language may employ more than one type of syllable structure, and it is to this that I now turn.

4.5
Turkish Syllable Structure

Turkish is a language which, in GP terms, licenses domain final empty nuclei as illustrated in (58). The two sets of words in (58) differ in their behaviour following the addition of suffixes. In (58a) the length of the final full vowel of the stem remains short in all forms. In (58b), however, the stem-final vowel appears to be lengthened when a vowel-initial suffix is added, but otherwise remains short. These forms with vowel lengthening are originally Arabic and Persian loans, and have been traditionally analysed as undergoing a process of vowel shortening in closed syllables, i.e. the final vowels in the stem are said to be underlyingly long. The stem-final consonant is syllabified into the coda of the stem-final syllable in both the nominative singular and plural, and in the onset of the final syllable in the possessive.

(58)

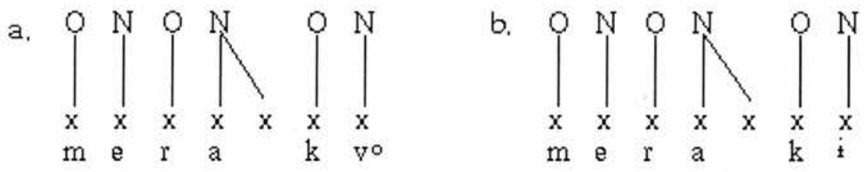
Nom. sg.	nom. pl.	3 sg. poss.	
a.			
kep	kepe	kepler	'cap'
sap	sapa	saplar	'stalk'
ahmet	ahmede	ahmetler	'Ahmed'
b.			
zaman	zama:nɨ	zamanlar	'time'
merak	mera:kɨ	meraklar	'law'
usul^y	usu:l^{y}ü	usul^{y}l^{y}er	'method'

As Kaye (1990) points out, the consonantal clusters created in the 3 singular possessive forms are generally analysed as either branching onsets or a coda-onset sequences. These are doubly anomalous in that Turkish does not otherwise contain branching onsets, and as coda-onset sequences they seem not to obey the normal phonotactic constraints of such sequences, e.g. we expect /l p/ but not */p l/. These anomalies, suggests Kaye, argue against the traditional vowel shortening account. Kaye represents /merak/ as (59), and /mera:kɨ/ as (59b), with the /k/ followed by an empty nucleus in both cases.[35] Kaye notes that given the syllabifications in (59), where the final vowel of the stem is underlyingly long, shortening takes place when the following nucleus is a licensed empty nucleus. Shortening before a licensed empty nucleus then becomes a general parameter which is switched on in Turkish. In (59a) the nominative singular ends in a licensed empty nucleus so the preceding vowel shortens, while the following nucleus in the possessive is not licensed, allowing the vowel to surface as long. The main criticism of this approach to vowel shortening is simply that there is no non-arbitrary connection between the shortening

[35]Kaye analyses the possessive suffix as {v°nv°}, the /n/ deleting in certain circumstances. The final empty nucleus fails to license the preceding empty nucleus, which is then realised, a necessary approach as the cold vowel /ɨ/ is part of the underlying inventory of Turkish. If the possessive suffix were simply /ɨ/ it would not surface, remaining phonetically inaudible. As I do not identify empty vowels with any particular segment comparable to the cold vowel of GP this problem does not arise.

of the vowel and the licensing or otherwise of the following empty nucleus. Kaye's analysis, while correctly describing the environment, lacks explanatory adequacy.

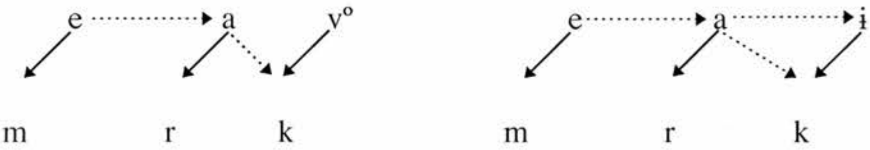
(59)



If coordination takes in Turkish the same form as in Italian with syllabification running from left to right so that onset consonants do not form an indirect government relationship with preceding syllables, then the AP representation of /merak/ is that in (60a), with /mera:kɨ/ in (60b). Again unlike Icelandic, vowels must coordinate with a following vowel if one is present, so that in /mera:kɨ/ in (60b) /e/ coordinates with /a/ which in turn coordinates with /ɨ/. This is supported by the existence of extensive vowel harmony in Turkish, this vowel harmony going in the same direction as the vowel to vowel coordination, i.e. from left to right. In /mera:kɨ/ the vowel in the penultimate syllable is also long, hence its coordination with both the following vowel and consonant. The same vowel to vowel coordination is demanded as in Italian, that a vowel will coordinate with a following vowel if one is present and has physical content, and if no vowel to vowel coordination is possible then coordination will be with a following onset. If nothing follows then no coordination is necessary. This, of course, is precisely the demand placed upon onset heads in Icelandic but in reverse. In (60a) the following vowel is the empty vowel, with the result that the /a/ coordinates with the following /k/ and the vowel remains short. The coordination necessary to produce long vowels in Turkish is thus equally a function of consonants as it is of vowels, the basing of vowel length on a combination of vowel-vowel and vowel-consonant relationships avoiding the arbitrary nature of lengthening/shortening of GP while preserving GP's insights as to the environments which condition vowel length.

(60)

a) b)



The words containing long vowels and consonants given above are, as already noted, originally loan words though long since assimilated to the Turkish phonological system. There exists, however, another set of words containing long vowels, derived generally from native Turkish words which differ greatly in their behaviour from the merak-type vowels. These forms have posed a number of problems and have been widely discussed (e.g. Clements & Keyser 1983; Kornfilt 1986; Sezer 1986). In addition there is a third set of words which also contain long vowels which behave in a still different way. The nominative of all the forms in (61) (taken from Clements & Keyser 1983) shows long vowels word-finally, contradicting the analysis given above where vowel length was said to be dependent on the presence of both a following consonant and a following vowel. If the analysis given here of merak-type words is correct, then we must somehow account for the forms in (61) in some other way.

(61)

nom.	**nom. pl.**	**dat.**	**2 pl. poss.**
a.			
da:	da:lar	daa	daɨnɨz
čɨ:	čɨ:lar	čɨa	čɨɨnɨz
či:	či:ler	čie	čiiniz
b.			
la:	la:lar	la:ya	la:nɨz
imla:	imla:lar	imla:ya	imla:nɨz
bina:	bina:lar	bina:ya	bina:nɨz

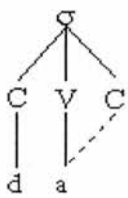
These 'dağ-type' words (after the orthographic form of the first word 'dağ' /da:/ in (61a)) historically contained a voiced velar fricative (represented in the spelling system by 'ğ' and known as yamusak ğe), and still do in many, particularly eastern varieties of Turkish, but this has been generally deleted in the modern standard language. The fricative has left its mark, however, in that in certain circumstances these words seem to behave as if they still contain a word-final consonant of some kind. This behaviour is reflected in the dative forms of (61). The dative suffix, as can be seen in (61b), begins with a /j/ which is lost when it is attached to a word ending in a consonant e.g kepe, *kepye, dative singular of kep 'cap'. The /j/ is retained following the final vowels in (61b) but lost following the ostensibly word-final vowels in the dağ-type words of (61a), apparently confirming that they are phonologically consonant-final, a fact reflected in a number of other ways. For example, the 2 plural possessive suffix has an initial vowel which is deleted after a stem-final vowel, e.g. dereniz, *dereɨniz, 2 pl. poss. of dere 'river', but we do not find this deletion following dağ-type words. In contrast the words in (61b) behave in all ways as if they were vowel-final and maintain therefore vowel length in all forms, ruling out any grouping of V:$ and VC$ syllables as part of a single process.

This aspect of the forms in (61a) led Clements & Keyser to propose that dağ-type words have the underlying form in (62a), with stem-final empty consonant slots to which the vowel will spread in certain circumstances, and which triggers (or blocks in the case of vowel deletion) the deletion processes mentioned above. The dative is derived as in (62b), reflecting the bisyllabic nature of /daa/ as opposed to the monosyllable of /da:/. These both contrast with the structure of /la:/ where the vowel is always doubly linked to two vowel slots and thus never varies.[36] In simple terms, dağ-type words behave as if they were consonant-final because they in fact are.

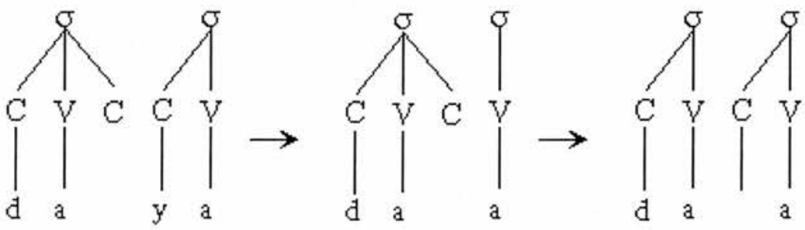
[36]Archangeli (1985) provides roughly the same analysis but without C and V labels and incorporating a minimal amount of hierarchical syllable structure, but the differences between the two are negligible.

(62)

a.



b.

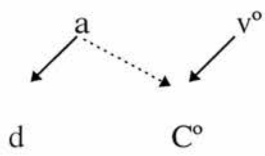


The simplest way within AP to handle the forms in (61a) is to claim that vowel length is not dependent upon a following consonant in these words but is instead derived in the same way as in Icelandic. Initially it might seem undesirable to have two separate kinds of long vowel within a single language (a point also made by Kornfilt (1986). However, there are clear physical arguments for distinguishing at least between merak- and dağ-type words. Rudin (1980) reports that the long vowel in dağ-type words is on average 13% longer than the long vowel in merak-type words, and is similar in duration to VCV sequences. If the two types of words do in fact differ on the surface then we would naturally expect them to differ underlyingly. However, I suggest that there is no third way of forming long vowels, so that la:- and dağ-type words pattern together, having essentially the same underlying structure, in opposition to merak-type words, as I shall attempt to show below.

If we assume, along with Clements & Keyser and Kornfilt (1986) that deletion of ğ left behind it an empty C- (or syllable-) slot, dağ would possibly have the representation in (63), with a final empty nucleus licensing an empty consonant. Empty consonants would behave in similar ways to empty vowels, so that for instance they could not indirectly govern, nor be indirectly governed by, another segment. Hence, from the viewpoint of the empty consonant in (63) there is no requirement for it to coordinate with either of the flanking vowels. However, as the

/a/ in (63) can not coordinate with the final empty nucleus it must instead coordinate with the consonant, and although the consonant itself is empty, coordination with /a/ would require it to have physical content. The result would be that the consonant would be audible e.g. /daɣ/ with both a short vowel and a final consonant, which is not the result we require.

(63)



The problems for an AP interpretation incorporating empty consonants are compounded by the apparent presence of empty consonantal (or syllable) slots word-internally as well as word-finally. Kornfilt (1986) suggests that the long vowels in the accusative forms in (64), all words spelt with yamusak ğe, are best described as arising from vowel spreading to empty consonantal slots created by deletion of the velar fricative. The forms without long vowels instead undergo a process of epenthesis (the epenthetic vowels are underlined). In the nominative and locative I assume that the stem-final consonant is licensed by an empty vowel. If we assume that the nominative forms in (64) have the structure CVğ$V^{o}CV^{o}$, the initial full vowel must coordinate with the empty consonant as noted above. If we assume that somehow it can avoid this coordination and instead coordinate with the following empty vowel, resulting in apparent epenthesis as desired, and the locative forms behave similarly. There is no way, however, given these forms, to derive the long vowels of the accusative.

(64)

	Nom	Locative	Accus	
a.	čɨɨr	čɨɨrda	čɨ:rɨ	'era'
b.	uur	uurda	u:ru	'good luck, good omen'
c.	baɨr	baɨrda	ba:rɨ	'rump'
d.	böür	böürda	bö:rü	'side, flank'
e.	oul	oulda	o:lu	'son'

Epenthesis is not confined to words which originally contained a velar fricative, as (65) shows (the epenthetic vowels again underlined). The nominative and locative of the words in (65) behave in the same way as the corresponding cases in (64), while the accusative shows the predicted consonant to consonant coordination. Kornfilt suggests that in the nominative and locative of all these forms syllabification following epenthesis is (C)V$CVC, while the accusatives syllabify as (C)VC$CVC. As noted above, Kornfilt assumes that in forms such as /u:ru/, with VğCV as its underlying form, resyllabification of ğ, i.e. the empty slot, into the first syllable, allows the initial vowel to spread.

(65)

	Nom	Locative	Accus	
a.	burun	burunda	burnu	'nose'
b.	ʃehir	ʃehirde	ʃehri	'city'
c.	akɨl	akɨlda	aklɨ	'intelligence'
d.	bahis	bahiste	bahsi	'bet'

Kornfilt is unable to account for the alternative forms in (66) (taken from Sezer 1986), however, where the differences between the forms in (64) seem to vanish.[37] Sezer (1986) notes that vowel assimilation occurs in certain words, involving lowering of the target vowel, the environment of which remains vague. Assimilation is optional and generally occurs in less formal speech, and in addition appears to be commoner in some words than in others, so that 'standard' forms such as /aɨz/ may actually be much less common than e.g. /aaz/, even in formal speech. We should also note that assimilation is not a casual speech process of the same type as those we have seen in earlier chapters, e.g. nasal assimilation in 'tem pence' but a phonological one with a clear lexical input. It seems clear however from Sezer's description that vowel assimilation only targets the epenthetic vowel. In addition, there is a process of monophthongisation, which Sezer describes as also optional though favoured in many cases, which turns a series of adjacent identical vowels, whether created by vowel assimilation or underlying, into single long vowels. Sezer assumes that the epenthetic vowel in the forms in (64) is in fact underlying and that there is no empty slot. The alternative forms in (66) are then somehow created by vowel assimilation and/or monophthongisation, and Sezer further suggests that the 'normal' accusative forms with long vowels in (65) are also created by the same processes.

(66)

	Nom	**Accus**	
a.	čɨɨr	čɨ:rɨ / čɨɨrɨ	'era'
b.	uur	u:ru / uuru	'good luck,
c.	baɨr / ba:r	ba:rɨ / baɨrɨ / baarɨ	'rump'
d.	böur / böör / bö:r	bö:rü / böürü / böörü	'side, flank'
e.	oul / ool / o:l	o:lu / oulu / oolu	'son'
f.	aɨz / aaz / a:z	a:zɨ / aɨzɨ / aazɨ	'mouth'

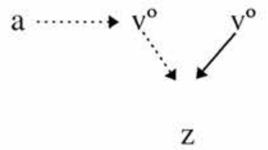
Sezer's approach fails on two levels. It fails to adequately account for the environments in which assimilation and monophthongisation take place, as he himself acknowledges. It also fails to both account for the lengthening in dağ-type

[37]Although Sezer gives data for the 3sg possessive and not the accusative, the forms are in fact

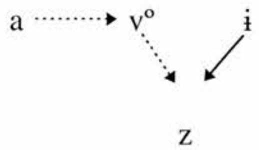
words and to provide a link between these and other forms in which yumasak ğe has been lost word-internally. I suggest, however, that we adopt Sezer's analysis of the underlying forms free of empty slots but with the epenthetic vowel already present. In (67) I give the underlying AP representations for both /aɨz/ and /a:zɨ/, shorn of any but the most basic governing relationships. The two differ only minimally, the latter terminating in a full vowel, the former with an empty vowel. All that remains is to derive the correct surface realisations.

(67)

a)



b)



In neither form in (67) is there an empty slot corresponding to 'ğ', and I assume that deletion of the velar fricative in modern standard Turkish is complete. The initial /a/ in both words must be integrated into the prosodic hierarchy by coordinating with the following vowel, and although this vowel is empty it must now receive physical content, causing it to be realised as /ɨ/. The resulting nominative form predicted by (67a) is /aɨz/, which indeed is the standard form. While Turkish, as noted already, has an obligatory process of vowel harmony, the alternative realisations /aaz/ and /a:z/ for the nominative strongly suggest, as Sezer observed, the existence of optional processes of vowel assimilation and monophongisation. Vowel assimilation occurs in environments such as (67a) where one headed nucleus indirectly governs a following headed nucleus with no intervening consonant, so that the accusative following vowel assimilation will be /aaz/ where the empty vowel is realised as a copy of the full vowel. In addition, (67a) appears to be being reinterpreted in modern Turkish as if the relationship between the two initial vowels is one of both direct and indirect government, i.e. as if it were a single long vowel of the same kind as that already present in words such as /la:/ where, as in Icelandic, the

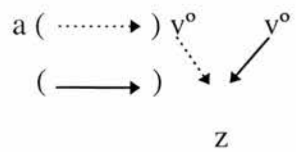
---

identical and I refer to them throughout the following discussion as accusative forms.

/a/ is licensed to govern a dependent vowel which is identical to itself. In both cases the vowel copying is an alternative to realising the empty vowel as /ɨ/, and in both cases it is optional, depending on both the formality of the speech context and in some cases on the particular word.

We can represent these processes as in (68), where indirect licensing between the initial vowels will always be present, but direct licensing may not be. In the nominative, assimilation and monophongisation will optionally occur, though becoming commoner in less formal speech. Assimilation and monophongisation thus result in the nominative becoming identical to the accusative forms with long vowels, but the accusative itself may also go in the other direction and become identical with the nominative (aside from the presence of final /ɨ/) and show realisations with hiatus sequences of identical vowels or of short vowel plus /ɨ/. This suggests that in informal speech, the nominative and accusative are not distinguished. In formal speech, however, assimilation and monophongisation (i.e. the presence of direct and indirect licensing between the vowels) are triggered by the accusative alone, resulting in a distinction between /aɨz/ for the nominative and /a:z/ for the accusative. However, it seems clear that the formal distinctions between the nominative and accusative are fast disappearing, with assimilation and monophongisation gradually dominating in the modern language.

(68)

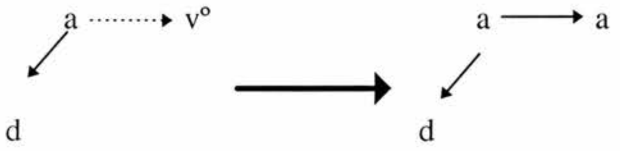


/aɨzɨ/ ~ /aaz/ ~ /a:z/

Exactly the same processes must be at work in dağ-type words as in (69). The underlying structure is again CVV°, the velar fricative being deleted and leaving no trace, though unlike the other forms with yumasak ğe the vowel-vowel sequence is here word-final. The initial /a/ indirectly governs the following empty vowel, but this only gives us the form */daɨ/. I suggest that while assimilation and monophongisation in words with yumasak ğe is spreading, in these words where the velar fricative was word-final these processes are now obligatory and lexicalised. In the accusative, the accusative suffix replaces the empty vowel with a full vowel which the initial /a/ is unable to directly govern, resulting in /daa/. We can thus see that vowel length in these words has the same basic structure as in /la:/, although it is derived somewhat differently. The long vowels in both of these groups are clearly phonologically distinguished from those in merak-type words, and this is further reflected in their phonetic realisation. There is no need for resyllabification of any kind, or for compensatory lengthening, nor for any processes of epenthesis particular to Turkish.

(69)

a)



4.6

Summary

The structures described in this chapter meet the goal of describing phonology solely in terms of the relationships between gestures, in particular syllable structure and segmental length are represented without recourse to any external timing slots of any kind. Building on the types of relationships between gestures found within segments, syllable structure can be formed from simple binary head-dependent relationships between segments, some of them universal, others language particular. These abstract phonological relationships directly describe the physical relationships involved, and any changes in these phonological relationships will therefore have a direct effect on the physical output, as we would expect.

AP recognises three basic constituents, the nucleus, onset and coda. Each of these constituents contains at least a head and possibly a dependant, the head and dependant sharing a direct government relationship. Segments are incorporated into the prosodic hierarchy through indirect licensing, which is achieved by entering into an indirect government relation with another segment. Consonants in all languages must be indirectly licensed by a vowel, though as we have seen they may also enter into indirect government relationships with other consonants, whether within an onset, onset to coda, or, in the case of syncope, onset to onset. Nuclear heads, on the other hand, are not required to enter indirect government relationships in languages such as Icelandic, but in contrast must enter such relationships, I argue, in languages such as Italian and Turkish. These languages fundamentally differ both in their syllabification, right to left in Icelandic and left to right in Italian and Turkish, and in the construction of vowel length.

The phonological relationships put forward severely constrain the types of gestural coordination we expect to find, but at the same time the phonological relationships are also dependent upon the coordinations possible between gestures. While the data analysed here suggests that AP is capable of describing complex phonological phenomena at both the segmental and syllabic level, the range of data covered remains small and thus the structures proposed can only be taken as preliminary. In the same way, only a subset of the possible types of gestural coordination have been examined, and there may be a number of possibilities which should be part of any overall syllable structure. A more detailed study of the languages discussed here within AP would perhaps necessitate some changes both in the particular analyses put forward and in the general structures.

Chapter 5

Predictions and Problems

5

Introduction

Gestures, as real physical objects with spatial and temporal dimensions, present a potentially rich source for interpreting the complex structures of natural languages. In order to make progress in this area, however, there is a clear need to establish a non-arbitrary and constrained set of phonological relationships which will allow us to use gestures as a model for phonological processes. The task of this thesis is to build such structures at two levels, the level of internal segment structures and the level of intersegmental relationships. In addition, we are further constrained by the need to base such structures directly on the relationships which hold between gestures, so that there are no equivalents of root nodes, syllable nodes, moras, and a range of other objects which are common in some form to many phonological theories. Instead, gestures are the sole constituents of AP, with segments and syllables derived directly from the gestures and the relationships between them.

Although the set of data covered in the preceding chapters is relatively small, the resulting proposed phonological relationships are, of course, applicable cross-linguistically and subject to empirical testing. Whether or not the results are confirmed is a matter for further research. The set of phonological relationships put forward is also relatively small, and will certainly need augmenting in a number of respects. Looking, for example, at syncope in Icelandic, one would assume that some equivalent of the foot plays a significant role, and the incorporation of this into AP would necessitate a new set of phonological relationships. An analysis of all such areas is beyond the scope of this work, but there are a number of problems arising directly from the structures already proposed which reflect, to some extent, some of the apparent underlying problems of a gestural approach. These areas have been raised by a number of commentators, questioning either supposedly inherent advantages of AP over other non-linear phonological theories, or areas where the very structure of gestures rendered them incapable of providing a satisfactory description of certain data.

In this chapter I wish to discuss two of these areas, covering both the internal structure of segments and intersegmental relationships, which raise a number of important questions regarding the future development of AP. In particular I examine notions of weakening, both in terms of weakening of syllable structure and of segment-internal weakenings such as lenition. In section 5.1 I examine some implications of the syllable structure proposed so far, and in 5.2 and 5.3 I examine in more detail the processes of vowel epenthesis in Winnebago and Scottish Gaelic. I argue that these processes may be viewed as weakening of the phonological relationships between segments. In section 5.4 I examine the claims of Steriade (1990) and Kingston & Cohen (1992) that the relationship within gestures between constriction location and constriction degree should in certain circumstances be weakened, showing how apparent processes of spreading and spirantisation in a number of languages can be handled without recourse to such a weakening.

5.1
Intersegmental Coordination

Rather than relying on an arbitrary three way distinction of gestural overlap - minimal, partial and total - to account for the various types of segments found in natural languages, the theory of segmental structure set out in chapter three instead derived the physical relationships between the gestures in a segment from a small number of simple phonological relationships. Each segment consists of one or more gestures which may be a head or a non-head, with each segment consisting of at least one head. Each head then corresponds to a single event, so that in e.g. a postaspirated stop [p^h] both the labial and the glottal gestures are heads, and therefore there must be two corresponding events, resulting in a physical coordination in which the target of each gesture is offset with respect to the target of the other. The precise patterns of coordination within segments will naturally vary from language to language, but stable patterns are created wherein phonological relationships are directly and non-arbitrarily reflected in the physical realisations of segments. Moreover, the close phonological relationship between the gestures in e.g. [p^h], where each is part of a single segment, is reflected in the high degree of overlap which they exhibit.

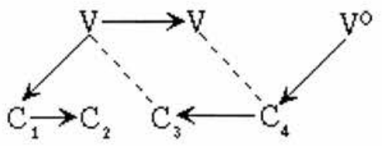
This connection between phonological relationships and physical coordination is continued into the domain of the syllable. (1) illustrates the various types of phonological structures suggested for Icelandic in chapter 4.[1] The first thing to note is that segments are arranged on two separate tiers, a vowel tier and a consonant tier, reflecting the gestural view that consonants are overlaid on vowels. Given the physical nature of AP, these tiers consist solely of gestures, that is, there are no timing units or syllable slots of any kind. Gestures form the basic units at all levels of the phonology. There is a dependency relationship between the two tiers, however, in that consonants are always dependent upon the prior existence of a licensing element on the vowel tier, so that while a consonant may be the head of its local domain, it is always ultimately directly licensed by an element on the vowel tier, whether by a full vowel or an empty one. Thus the two tiers are intimately connected. Given a CV sequence, the vowel directly licenses the consonant, and the two show a stable coordination whereby the midpoint of the consonant coordinates with a point early in the cycle of the vowel.[2] Although the two elements in a CV sequence belong to different domains, and therefore form separate constituents, the phonological relationship between them, with the consonant being governed by the vowel, can be read as creating a larger constituent which we can refer to as a CV-syllable.

---

[1] In future I will omit the arrows from the dotted lines representing indirect government, except where this may cause confusion.

[2] Whether or not this stable coordination, whatever its precise realisation, gives rise to the phonological relationship, or whether the reverse is true, is a moot point. If we were able to identify all of the stable physical relationships and determine whether each of them was phonologically relevant, then we could establish a direct causal link. Such a link is, however, yet to be determined, and the most we can claim at present is that the relationship between the phonological and physical relationships is non-arbitrary.

(1)



Although the consonant in a CV sequence is directly licensed by the vowel and dependent on it, both are heads of their respective domains. The headed status of such consonants allows them to directly license a dependent consonant, so that in (1) C_1 directly licenses C_2 and they together form a single constituent which we can refer to as an onset. The constituent nature of such a sequence is again derived from the governing relationship between them, though in this case C_2 is a non-head with C_1 the head of the domain. The physical relationship between the two consonants can be said to reflect the close phonological relationship between them, with the offset of C_1 coordinating with the target of C_2. On its own the connection between the phonological relationship and the physical realisation is not clear here, but it becomes so when we compare it with the coordination between C_3 and C_4 in (1). Here C_4 is the head of an onset, directly licensing C_3 to occupy the coda position. Unlike the onset dependent C_2 however, C_3 is the head of its domain, and as such the phonological relationship between the two consonants is not as close as that between the consonants in an onset.[3] As a result, the two show considerably less overlap, the offset of C_3 coordinating with the onset of C_4.

The different status of the two consonant clusters is also to be seen in the coordination they show with the segments on the vowel tier. Browman & Goldstein suggested the existence of a C-centre, basing their proposal on the behaviour of CCV sequences, where the consonants coordinated with the following vowel as if they formed a single unit. We can now see that this simply reflects the close phonological relationship between the consonants within an onset, where they form part of a single

[3] It seems clear that the separation of the rhyme into separate nucleus and coda constituents will necessitate much reanalysis of many languages. While such analysis is beyond the scope of this work, I argue that at least the data presented here supports this view. This does not presuppose that such a division may not, in fact, be true of all languages, as we have already seen in chapter 4 that it is not the case that all languages show the same relationships between segments.

domain, and their indirect licensing through the following nucleus. In contrast, although a coda-onset sequence C_3-C_4 also forms a single governing domain, the different status of the two consonants means that while C_4, as head of the onset, continues to coordinate in the expected fashion with its licensing vowel, C_3 does not coordinate with the following vowel at all.

As we have seen, the same general observations can be made for the segments on the vowel tier, that is, the closer the phonological relationship between two segments, the greater the amount of overlap. However, given the nature of gestures and of speech, phonological relationships and their physical correlates are not immutable, hence the existence of gradient phenomena such as the apparent deletion of the final /t/ of 'perfect' in 'perfect memory'. Over time, such processes can be realised as no longer gradient but categorical, so that the /t/ in 'perfect' may eventually come to be permanently lost (as it in fact has been in some dialects). Here the physical relationship would appear to precede the phonological, and this kind of shift doubtlessly lies behind e.g. syncope in Icelandic, or the hiatus vs long vowel contrasts of Turkish discussed in the previous chapter, i.e. originally casual speech processes caused by relatively small increases in the overlap between gestures, which ultimately become fixed.

Syncope, within the terms of the structures proposed here, is a process whereby two separate onsets which originally shared no phonological or physical relationship, come in time to have both. To take the example of 'b(e)ret' again, the /b/ and /r/ occupy separate onsets, but in casual speech they may each overlap the initial unstressed vowel to such an extent that they also overlap each other, causing that vowel to be hidden and giving rise in the process to apparent deletion of that vowel. All of the casual speech processes discussed so far have been of this type, where increasing overlap between gestures leads to a number of gradient phenomena. The same types of behaviour, as we have seen, also lie behind a number of categorical phonological processes. The assumption made in chapter 4 is that if two gestures or segments are directly physically coordinated, there must be a phonological relationship between them. In Icelandic 'Bitil' vs 'Bitli', I assume that syncope was originally a gradient process, as it doubtlessly still is in English 'beret', where in 'Bitli' the /t/ and /l/ would have increasingly overlapped the intervening vowel. This only becomes a categorical process once the overlap of the consonants is interpreted as involving a deliberate physical relationship between them, rather than an accidental one. Once this direct physical coordination is established, a phonological relationship must exist between the consonants. This leads us to speculate as to the possible existence of processes which are the reverse of syncope. In other words, what happens if two segments which overlap, such as the consonants in an onset, should, over time, drift apart to such an extent that they no longer overlap each other at all?

5.2
Vowel Epenthesis in Winnebago

Steriade (1990) suggested that such an explanation might lie behind the phenomenon known as Dorsey's Law in Winnebago, a member of the Siouan family. (Miner 1979) notes that a number of CVCV sequences in Winnebago correspond to CCV sequences in closely related languages such as Chiwere. The process is interpreted as one of epenthesis breaking up an underlying consonant cluster, where V^1 is a copy of V^2; C^1 consists of one of the obstruents /p k č s š x/, and C^2 of one of the sonorants /m n r w j/. The data in (2) are taken from both Steriade and Miner, with the putative epenthetic vowel in bold.

(2)

a. sh-wa-zhok	→	sh**a**wazhok	'you mash potatoes'
b. ho-sh-wa-zha	→	hosh**a**wazha	'you are sick'
c. hi-kro-ho	→	hik**o**roho	'he prepares'
d. hi-pres	→	hip**e**res	'know'

e. wakri-pras	→	wak**i**rip**a**ras	'flat bug'
g. wakri-pro-pro	→	wak**i**rip**o**rop**o**ro	'spherical bug'
h. knuʃ	→	k**ũ**nũšk**ũ**nuʃ	'cartilaginous'
i. kre	→	k**e**re	'he departed returning'

Miner cites a number of points which suggest that Dorsey's Law is in fact a synchronic process. Winnebago has a productive process of reduplification, and although CVCV sequences are not normally targets for reduplification, (2g,h) show that the sequences in question are. In fact, they are the only CVCV sequences which are so affected, suggesting that at some level they are actually CCV sequences, with epenthesis following reduplification. Winnebago also has a process of vowel nasalisation which affects the peripheral vowels /a i u/, nasalising them when they follow one of the nasals /m n/. In (2h) we see that if the following vowel is nasalised then so too is the epenthetic vowel, again suggesting that nasalisation occurs before epenthesis, which then copies the now nasalised vowel. Finally, there is a process lowering morpheme-final /e/ to /a/ before certain suffixes, e.g. /mããče/ 'he cut a piece off' + (h)ire becomes /mããčaire/ 'they cut a piece off'. Similarly we find e.g. /k**e**re/ 'he departed returning' becoming /k**a**raire/ 'they departed returning', where the vowel lowering must precede epenthesis.

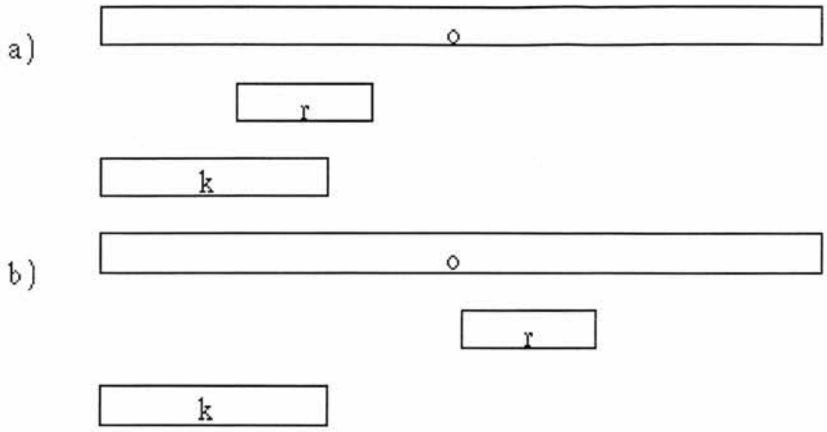
Dorsey's Law, then, seems clearly to be a synchronic process of vowel epenthesis into certain CCV clusters, where the epenthetic vowel is a copy of the original vowel. While epenthesis is not as such a rare process, there are a number of features of these epenthetic CVCV clusters which are, to say the least, unusual. Miner notes that these CVCV sequences are recognised as being unusual by native speakers, who refer to them as 'fast sequences', the name deriving from the fact that they are spoken, and apparently sung, faster than ordinary CVCV sequences.[4] In other words, fast sequences are shorter in duration than other comparable CVCV sequences. In addition, Miner notes that the accent (consisting of higher pitch and increased intensity) normally falls on every third mora, or if no third mora is

[4]Miner notes that while Chiwere has no similar fast sequences, other Siouan languages do

available it falls on the second mora. In words such as /kãnãk/ this should leave the first syllable unaccented, but in fact the first syllable may have nearly as much, or even as much, accent as the following. The generalisation which seems to arise from these facts is that despite the appearance of two full surface vowels in fast sequences, they nevertheless behave phonologically as if they were the manifestation of a single vowel.

Steriade proposes a novel interpretation of these various facts, one which preserves the underlying monosyllabicity of fast sequences by providing a gradient interpretation of syllable structure. Steriade represents the change from /kro/ to /koro/ as in (3). In (3a) the consonants overlap each other, and in turn they overlap the accompanying vowel, and this can be regarded either as the underlying syllabification or as an earlier stage in a diachronic process. The important point to note is that the vowel gesture extends throughout the period when the consonants are active, so that the initial part of the vowel is hidden. Steriade then suggests that there is a change in the coordination of the various segments, so that the sonorant /r/ moves rightwards, coordinating with a point later in the vowel. The result will be to uncover that portion of the vowel which was hidden by the combination of obstruent and sonorant, giving the impression of epenthesis while in fact maintaining the underlying monosyllabic structure. In other words, fast sequences are in fact monosyllables. This would remove the need for any processes of vowel insertion or copying, and would provide a straightforward explanation for the nasalisation of all vowels in e.g. /kũnũškũnũš/ and for the vowel lowering in /karaire/: only a single vowel is affected in each case.

(3)



Such an approach has other obvious attractions. The fact that fast sequences are accented more like monosyllables than like bisyllables is not surprising if Steriade's interpretation is correct, nor is the shorter duration of fast sequences compared to ordinary CVCV sequences. As long as the vocalic gesture itself were sufficiently large, the shifting of the sonorant rightwards would be able to reveal a portion of the vowel's target which would be identifiable with a full vowel. The overall duration of the two vowels would still be close to that of normal short vowels, if not identical, hence the lesser duration of fast sequences compared to other CVCV sequences. What then if the sonorant were to be moved rightwards, again failing to overlap the preceding obstruent, but to a much lesser degree? Steriade suggests that if the rightwards movements were of a significantly lesser degree, the amount of 'hidden' vowel uncovered would be identifiable only as schwa. This may be a precursor of the type of epenthesis found in Winnebago, and Steriade cites the contrasts in (4) from Latin as evidence to support this.

(4)

patri pat**e**ri pat**i**ri

She suggests that the inserted vowel (in bold) in Late Latin 'pat**e**ri' is an indeterminate vowel, and that it results from a small rightwards movement of the sonorant. Presumably the sonorant in this and similar words would be realised shortly after the preceding consonant, thus coordinating with a point early in the vowel. Eventually, however, the movement may be greater so that the sonorant will coordinate with a point significantly later in the vowel, thus uncovering a greater portion of the following vowel, ultimately leading to forms such as 'patiri', so that the form with a full vowel simply results from a greater rightwards movement by the sonorant.

There are of course a number of obvious disadvantages with Steriade's analysis, as we have seen, primarily connected with what view of syllable structure, if any, is implied by the structures in (3). The greatest problem arises from the fact that (3) implies the existence of at least three separate, discrete areas of a vowel with which a consonant in the onset is free to coordinate. Taking a CCV obstruent-sonorant-vowel sequence, there must be a stable pattern of coordination both between the consonants themselves, and between the consonants and the vowel. Gradient movement rightwards of the sonorant might reveal a schwa-like portion of the vowel, but in order to do so it must still coordinate with the vowel, creating a second coordinative pattern between the two, and the phasing behind this coordination would need to be stable in order to consistently result in epenthesis (for want of a better word). Further, the sonorant is free to move still further rightwards, revealing a more substantial portion of the vowel, and in the process it must coordinate with yet another portion of the vowel, and again the phasing would need to be stable. Given this, there seems nothing to prevent a single language from employing all three types of coordination, resulting in a three-way CCV ~ CəCV ~ CV_iCV_i contrast carried only by a distinction in the coordination between the sonorant and the vowel. All three of these would be monosyllables. Add to this any possible alteration in e.g. the coordination between the obstruent and the vowel, and we once again generate too many different types of distinctive coordinative patterns.[5]

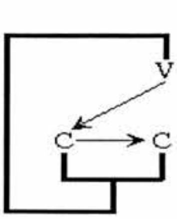
[5]Steriade suggests that the only categorically distinct processes are whether the shifting sonorant moves to a peripheral or a non-peripheral position. In Early Latin we find contrasts such as 'trapezita' ~ 'tarpezita' (table), where the sonorant /r/ can be viewed as moving to a peripheral position in the latter, as opposed to the non-peripheral target of the sonorant in 'patiri'. This still fails to distinguish between possibilities with a full vowel or with schwa..

Looking again at the relationships between the segments in (1) above, we can see a number of sites which have been frequently reported as sites of phonological weakening. In particular, segmental weakening is frequently cited in the coda and within the onset, and in English we have as examples of this the historical deletion of coda /l/ in words such as 'walk', and of onset /n/ in words such as 'knight'. These are both examples of segmental weakening as a result of their structural position, but there is another type of weakening which we should predict in AP, one which derives directly from the physical nature of gestures.

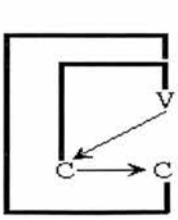
We recall that the existence of a physical relationship between two segments always implies the presence of an indirect government relationship. The relationship between the segments in (3a) can be represented as in (5a). Recapping somewhat, the phonological relationship between the consonants is realised by coordinating the offset of the first segment with the target of the second, and this is represented here by the bold line connecting the two. The relationship between the consonants constitutes a governing domain, and we can interpret it as a single constituent, an onset. Both consonants are indirectly licensed by the following nucleus, and the coordination of the consonants with the following vowel is also shown by a bold line. This, then, is an example of a set of phonological relationships being directly represented in terms of physical coordination. The vowel directly licenses the head of the onset, both segments being heads of their domain and the onset head in turn directly licenses a dependant. What would be the result, then, of a weakening in the power of the onset head to indirectly govern certain coda dependants? In other words, what happens if the obstruents in fast sequences are able to directly license sonorants as dependants in the onset, but are unable to indirectly license them?

(5)

a)



b)



We can see the result of such a weakening in (5b). There is no change at all in the direct government between the onset segments, only in the indirect government, the two consonants still forming a branching onset. However, the failure of indirect government to occur entails the removal of the physical relationship between them so that they no longer coordinate, i.e. they no longer overlap. While the coordination between the consonants fails, the same is not true of the relationship between the consonants and the following vowel. In other words both consonants, while not coordinating with each other, must nevertheless continue to coordinate with the vowel which indirectly governs both of them. Coordination between the head consonant and the vowel is the same as that in a CV sequence. As for the non-head consonant, C^2, I follow Steriade in claiming that the precise coordination may vary, resulting in the different phenomena which she noted. In Winnebago the sonorant coordinates with a point significantly late in the vowel so that enough of the vowel is uncovered for it to be identified not as schwa but as a copy of the following vowel. Crucially, however, these different types of coordination are in no way phonologically distinct. The sole distinction lies in whether or not the consonants overlap, giving a two way distinction of either CCV or CVCV, the identity of the epenthetic vowel varying as Steriade noted but with no possibility of a phonological contrast between a schwa or vowel copying.

The predictions of the structures developed here are borne out, and at the same time we are able to give a constrained account of the processes noted by Steriade while avoiding the problems discussed above. Fast sequences in Winnebago can thus be said to be monosyllabic, obviating the need either for a process of epenthesis or for a gradient analysis of syllable structure. All that need be said is that the relevant obstruents can directly license certain sonorants as dependants in the onset in the normal way, but that indirect licensing, and hence physical coordination, fails. What was almost certainly an originally gradient process, a process which might be described as syncope in reverse in which the consonants in the onset gradually drifted apart, has simply become categorical.

The indirect government relationship between onset segments can fail, of course, because the only absolute requirement for consonants is that they form an indirect government relationship with a vowel. Thus, while indirect government between onset head and dependant is the norm, as long as both consonants are indirectly governed by a vowel then indirect government between the consonants is not absolutely necessary. This naturally brings to mind the other common relationship which holds between consonants, that between codas and their following onsets. Coda-onset is another common site for processes of segmental weakening (Harris 1992), and given our analysis of Winnebago, it seems likely that a similar pattern should be found here, where coda and onset consonants drift apart. The absence of any evidence for such a process would seriously undermine the argument developed here. However, epenthesis in both Scottish and Irish Gaelic suggests that such a process does indeed exist.

5.3
Vowel Epenthesis in Gaelic

The data in (6) illustrate the situation in present-day Irish. Epenthesis acts to separate clusters of sonorants followed by any of a number of non-homorganic consonants (the range of consonants represented by C^2 varies according to dialect) in a similar fashion to Winnebago, except that the clusters in question here are coda-onset clusters. As for Winnebago, while we can assume that at some earlier stage of the language these forms contained no such vowels, all present-day dialects show an epenthetic vowel in these words which remains phonologically distinct in a number of ways from other underlying vowels.

(6)

a.

borb	[borəb]	'savage'
arm	[arəm]	'army'
dearmad	[d'arəməd]	'mistake'

fearg	[f'arəg]	'anger' (nom.)
feirge	[f'ar'ig'ə]	'anger' (gen.)
dorcha	[dorəxə]	'dark'

b.

téarma	[t'e:rmə]	'term'
léargas	[l'e:rgəs]	'insight'

There are a number of points of interest here. While C^2 is generally voiced, it may be a voiceless fricative as in 'dorcha'. However, if C^2 is a voiceless stop as in 'corp' (body) epenthesis does not occur so that the identity of C^2 is crucial in some yet to be defined way. Epenthesis is also blocked if the preceding stressed syllable contains a long vowel, as (6b) shows. As epenthesis thus only occurs following a short stressed vowel,[6] it is generally assumed that at least historically the sonorant was originally in the coda of the stressed syllable, while following epenthesis it is, or was, resyllabified into the onset of the syllable headed by the epenthetic vowel. The fact that the consonantal environment in itself is not sufficient to trigger epenthesis clearly implies that syllable structure has some part to play.

In contrast to the inserted vowel in fast sequences, the epenthetic vowel in Irish is always (a variant of) /ə/ or /i/, depending on the quality of the flanking consonants, and never a copy of the preceding vowel, and on the surface epenthetic syllables are not phonetically distinct from underlying syllables containing the same vowels. They remain phonologically distinct in two ways however. The forms in (6a) are entirely predictable in that they never alternate with forms where the epenthetic vowel is absent, and are thus never targeted by syncope processes which affect identical underlying unstressed syllables in similar environments. More importantly, a process of palatalisation affects stem-final consonants, whether single consonants or clusters, in environments such as the genitive singular of many nouns.

[6]Given that regular stress falls almost invariably on the initial syllable, epenthesis in (1) always occurs following primary stress. Epenthesis also occurs in phrasal environments in some dialects, so that e.g.

This gives a contrast between e.g. nom. sg. 'bàrc' [ba:rk] (ship) and gen. sg. 'bàirc' [ba:rʲkʲ], where both final consonants are palatalised. However, if the consonants are separated by a vowel only the final consonant will be palatalised, e.g. nom. sg. 'balóg' [bało:g] (ruin) and gen. sg. 'balóige' [bało:gʲə] where only the final stop is palatalised. In words with epenthetic vowels, palatalisation affects both the sonorant and the following consonant, as in the genitive form [f'ar'ig'ə] in (6a) where normally we expect only the stop would be targeted. Clearly while on the surface the sonorant and following stop are non-adjacent in epenthetic words, this is not the case at some point in the derivation.

Whether or not the aberrant behaviour of epenthetic syllables is merely lexical or due to an active phonological process is somewhat of a moot point. While dialect descriptions generally pay little attention to the synchronic nature of epenthesis other than to note its existence and list its environments, elsewhere discussion has tended to focus on the diachronic development, though not necessarily to the exclusion of the synchronic position. Phonological descriptions have usually seen epenthesis as part of a more general process of preservation, or redistribution, of syllable length. Ó Baoill (1980) suggests that at the time epenthesis originated a word such as 'arm' would have consisted of a sequence of four moras, one for the initial vowel, one for /r/ and two for the final /m/ which he suggests was geminated in Old Irish. Subsequent degemination of the /m/ led to epenthesis to the delinked mora and resyllabification, thereby preserving the overall mora count.[7] More recently Ní Chiosáin (1991) has suggested a similar process without the need for a bisyllabic /m/. She suggests that the sonorants in (6a) are assigned a mora by applying Weight-by-Position (Hyman 1985). Following Weight-by-Position, 'arm' will contain two moras, one attached to the vowel, another attached to the sonorant.

'an-mhaith' (very good) can be realised as [anə 'wa] where epenthesis occurs preceding the primary stressed syllable.

[7]Ó Dochartaigh (1981) takes a somewhat similar approach within the framework of Dependency Phonology, although arguing instead that it is the /r/ which must be geminated. Vocalisation of the first half of the geminate as part of a general process of 'vowel strengthening' i.e. increasing sonority, followed by resyllabification, again leads to /arəm/. While there is much disagreement as to whether either of the consonants in 'arm' were ever geminates, the importance of these approaches lies in their interpretation of epenthesis as a phonological process.

A general process of mora delinking then applies, delinking the sonorant from its mora. A special process of Mora Preservation then comes into effect, preventing deletion of the unlinked mora and causing epenthesis of a vowel to it, and subsequent resyllabification then leads to the correct surface form.

These and other analyses have much in common, such as delinking and relinking of features, abstract timing units, insertion of segmental material not previously present, and common to all is some degree of resyllabification. Such processes, as we have seen, are not part of the armoury of AP. Whatever the virtues of these different approaches, however, none of them are able to describe the more complex patterns found in Scottish Gaelic, patterns which echo those of Winnebago in a number of ways.[8] As is apparent in (7) epenthesis occurs in Scottish Gaelic in much the same environment as in Irish, though the range of consonants involved is generally larger. As in Irish, epenthesis is blocked if the preceding vowel is long or a diphthong.

(7)

a.

borb[9]	[bɔrɔb]	'savage'
arm	[aram]	'army'
dearmad	[d'ɛraməd]	'forgetfulness'
fearg	[fɛrag]	'anger' (nom.)
feirge	[fer'ig'ə]	'anger' (gen.)
dorcha	[dɔrɔxə]	'dark'
marbh	[marav]	'dead'

[8]Not all Scottish dialects have the type of epenthesis shown in (2), indeed in dialects such as East Sutherland (Dorian 1978) epenthesis is to all intents and purposes identical to that of Irish. However, as most of the recent phonological literature has concentrated on the Gaelic of the Outer Hebrides, and on the dialect of Barra in particular, unless otherwise specified it is these dialects which are under discussion here.

[9]Historical voiced stops have become devoiced in most dialects. The broad transcription here with voiced stops is given for ease of comparison with the Irish data. Similarly, neither post- nor preaspiration is marked. Nothing crucial hinges on this.

b.

mìorbhail	[miərvalʹ]	'miracle'
corp	[kɔrp]	'body'

The first point of departure from Irish in (7) is that the epenthetic vowel is generally a copy of the preceding vowel, with certain provisos. In southern Hebridean dialects such as that of Barra the [back] value of the epenthetic vowel agrees with that of the intervening sonorant, though this becomes less so the further north one travels so that in the most northerly dialect of Lewis the epenthetic vowel is almost invariably a copy of the preceding vowel (Borgstrøm 1940). Gaelic shows a much reduced range of vowels outside main stress, which in all but a very few words means anything other than the initial syllable, but epenthetic vowels show the same range of contrasts as do normal stressed syllables.

Leaving aside the question of the feature spreading, no model yet proposed has adequately accounted for the unusual prosodic features of the epenthetic vowel. Aside from a few borrowed words and some compound forms, stress falls on the initial syllable, other syllables being weakly stressed or unstressed. A form such as 'aran' (bread) thus has strong stress on the first syllable and weak stress on the second. We can compare this with 'arm' where the epenthetic vowel, rather than being unstressed, has stress at least equal to that of the preceding vowel, and Oftedal (1956) notes that if any syllable tends to have the greater stress in such a form it is the epenthetic. In addition, in all Hebridean dialects epenthetic words have a distinctive tonal structure as illustrated in (8), adapted from Oftedal (1956) for the dialect of Leurbost, Lewis. In the disyllabic forms of (8b) (syllable boundary in words with hiatus indicated by a hyphen) there is a pitch rise on the initial vowel and a fall at the onset of the second, similar patterns being found in the other dialects. Compare this with the forms in (8a). In each of these pitch rises on the initial vowel in the same way as in (8b) but rather than falling at onset of the second it continues to rise so that the long vowel in [bo:] shows a continuous rise in pitch throughout the vowel.

(8)

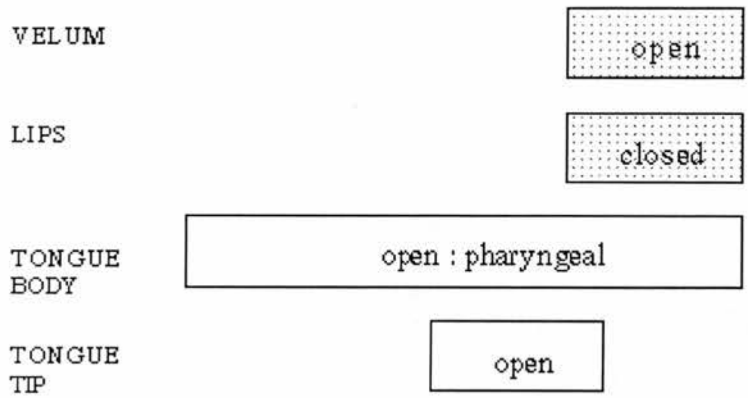
a.		b.	
b o o	'cow'	b o - o	'underwater ridge'
f i ə x	'debt'	f i - ə x	'crow'
a r a m	'army'	a r a n	'bread'

This provides us with a clear distinction between [aram] on the one hand and [aran] on the other. The crucial point to note is that the two vowels of [aram] have the tonal and stress patterns of a single long vowel or diphthong, and this is reflected in native speaker intuitions that [aram] contains either a monosyllable or something close to it. It seems difficult, if not impossible, to account for these facts as resulting from vowel insertion coupled with feature spreading from the preceding vowel. The simplest account would seem to be one in which the apparent identity between diphthongs, long vowels and epenthetic sequences is trivial - in other words they are identical because they are the same object. Borgstrøm (1938) makes essentially this point in attempting to show that the original form of epenthesis was something close to the present form in Scottish Gaelic : (Borgstrøm 1938, p38) 'In certain groups of comparatively open and sonorous consonants as *-rw-*, *-lx-*, etc., there was an interval between the two articulations during which the tongue was for a moment in an intermediate and relatively open position. This interval was not part of any of the consonants; its nature was more vocalic than consonantal. Part of the vowel preceding the consonants could penetrate into this 'vocalic point'; the one vowel was divided into two parts, and the new vowel-part had as much stress as the other, since they were felt to be only one vowel, or at any rate one syllable.'

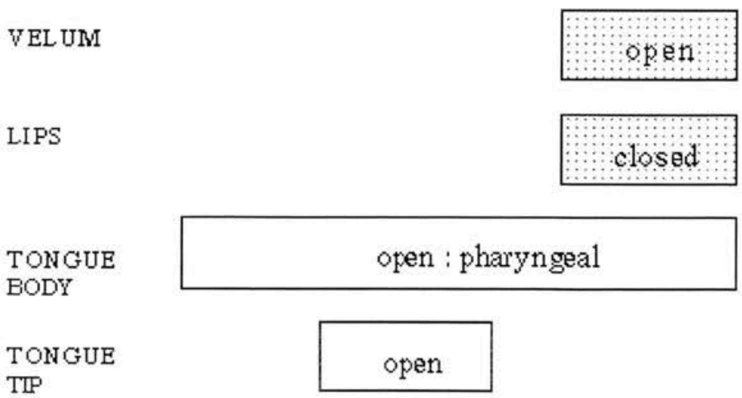
The parallels with Winnebago are clear, particularly in terms of the interaction of epenthesis and accent in Winnebago, and epenthesis, pitch and stress in Scottish Gaelic. Following Steriade's account of Dorsey's Law, Bosch (1995) has proposed a similar analysis of Scottish Gaelic which we can illustrate as in (9).[10] In (9a) we have a representation of the putative Old Irish pronunciation of 'arm', where the final consonants show the expected overlap. Epenthesis is then simply the result of the shifting leftwards of the sonorant, revealing the previously hidden vowel.

(9)

a.



b.



---

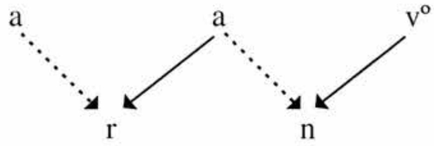
[10]By contrast, Bosch (1989) argued that the unmarked syllabification in Scottish Gaelic is VC$V, while epenthetic sequences are marked by syllabifying as V$CV. In this she follows both Clements (1986) and Kenstowicz & Kisseberth (1979), who in turn follow Borgstrøm (1940). Borgstrøm suggested a distinction in syllabification between, in his notation, *ar-an* and *m[a-ra]v*, where the sonorant in non-epenthetic words is syllabified leftwards, contrary to normal expectations, while that in epenthetic words is syllabified rightwards. In fact, other researchers (e.g. Oftedal 1956) have failed to notice this contrast, and there is little or no phonetic evidence to support it.

Unfortunately, this is subject to the same criticisms as Steriade's analysis, particularly in terms of identifying where it is that /r/ moves from, where it moves to, and how many other sites are available. In addition, in purely physical terms the earlier realisation of the sonorant would not in fact automatically provide us with the correct results. In the dialects of the Outer Hebrides, both the epenthetic and the preceding vowel are of normal duration (Borstrøm 1940, Ó Curnáin 1990). Moving the sonorant leftwards with respect to the vowel would automatically cover up much of the initial portion of the vowel, so without some additional process to somehow stretch it, epenthesis would automatically result in severe diminution and possibly deletion of the initial vowel. Further, the portion of the stressed vowel presumably underlyingly hidden by the target portion of the sonorant is mainly its offset, so uncovering it in a gradient as opposed to categorial fashion would again not necessarily lead to epenthesis of a full copy of the stressed vowel but more probably a heavily reduced version of it.

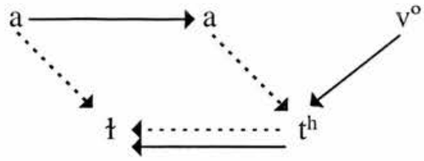
Again, what Bosch's theory lacks is simply a description of the phonological relationships between the segments in (9). The theory developed here provides us with the syllable structure in (10a) for a word such as 'aran' (bread), which has a surface structure consisting of two headed nuclei (ignoring the final empty vowel), i.e. two syllables. (10b), by contrast, shows the structure of 'alt' (joint), which on the surface consists of only one syllable. However, given the theory developed here, both words contain the same number and type of vowels (two instances of /a/ plus an empty vowel) but differ in the phonological relationships between them. In (10b) the coda /ɫ/ and onset /tʰ/ overlap the dependent /a/ of the first syllable in the expected way, leading to a surface VCC syllable. The stop in (10b) is able to directly and indirectly govern the preceding sonorant (shown by the double arrow notation) and thus coordinates with it.

(10)

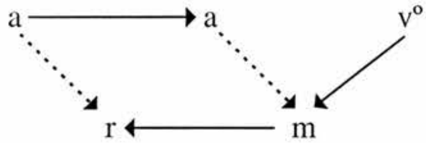
a) 'aran'



b) 'alt'



c)



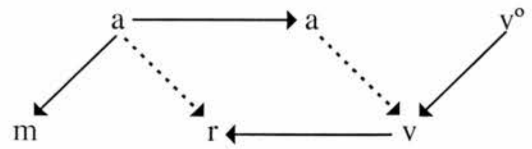
As in (9a) above we can assume that historically both 'alt' and 'arm' had the same structures, in both cases the sonorant being licensed to occupy a coda position by being directly governed by the following onset, /t^h/ in 'alt' and /m/ in 'arm' respectively. What both Borgstrøm and Bosch suggest is that for some reason the /r/ and /m/ in 'arm' fail to coordinate yet somehow remain part of the same syllable. Aside from the difference in the phonological relationships between the segments - onset head + onset dependant as opposed to coda + onset - the processes in Winnebago and Scottish Gaelic are identical. Epenthesis in both cases is simply a result of a reduction in the licensing ability of certain consonants which manifests itself in a failure of the relevant consonants to coordinate. In 'arm', as indicated in (10c), the /m/ directly governs the /r/ so that 'arm' has an apparently identical structure to 'alt' in (10b). They differ though in one crucial respect, that is while /m/ is able to directly govern the preceding /r/ it is unable to indirectly govern it and hence coordination between them fails. Not all onset segments are affected in this way, so that in Gaelic the less sonorous voiceless stops can coordinate with coda consonants in the normal way. The more sonorous the onset, the less licensing ability it possesses. Given our lack of knowledge as to the nature of sonority, and its manifestation in AP in particular, I leave the question of why it is that sonority in both Gaelic and Winnebago should equate with licensing ability to further research.

The failure of /r/ and /m/ to coordinate in (10c) leads to the uncovering of the dependent vowel of the stressed nucleus, resulting in its being audible and thus giving rise to apparent epenthesis. The fact that the epenthetic vowel is a copy of the preceding vowel is simply a reflection of the underlying identity of the dependent vowel. Perhaps most significantly there are no processes of resyllabification, a feature common to all previous descriptions, and thus no need for rule ordering of any kind. Again both of these are ruled out as possible processes in AP. Epenthesis involves only a single process which we can reduce to being simply part of the phonotactics, that is which onset segments can physically coordinate with which coda segments. No other 'rules' are involved. The sonorants in both (10b) and (10c) are in the coda of the initial syllable despite appearances to the contrary. This is seen more clearly in some southern dialects where historically intervocalic consonants, especially sonorants, are glottalised. In 'falamh' [faɫˀəv] (empty) the intervocalic sonorant is in an onset and glottalised, while that in 'folbh' [[fɔɫəv] (going), where the sonorant is in a coda, is not glottalised.

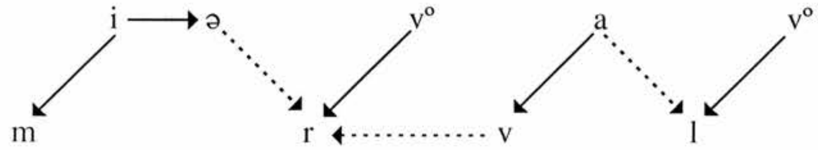
The forms in (11) show why epenthesis does not occur following surface long vowels or diphthongs. In (11a) 'marbh' has the same type of structure as 'arm' where the /r/ is syllabified into a coda by a following onset /v/. The /v/ being unable to coordinate with /r/ allows the dependent /a/ to surface as an epenthetic vowel. Compare this with 'mìorbhail' in (11b). Here epenthesis is not triggered simply because the required conditions are not present as the sonorant /r/ is not in a coda but in an onset. The relationship between the /r/ and /v/, again shown here by a solid connecting line, is of the expected onset to onset kind - /v/ is indirectly licensed by the following /a/ and thus does not coordinate with the preceding empty vowel but instead with the preceding onset, in this case /r/.

(11)

a) ‘marbh’



b) ‘mìorbhail’



The reason why long vowels, diphthongs and epenthetic sequences share common stress and intonation patterns is now clear - they are structurally identical. As noted above, for Lewis tone has been described as continually rising through a long vowel, and as showing the same continuous rise through a stressed vowel - epenthetic vowel sequence. This contrasts with the pattern in Barra where the stressed vowel shows rising pitch which then levels off, remaining at the same level throughout the epenthetic vowel, but in both cases the pattern is quite distinct from normal stressed vowel - unstressed vowel words.

Although Lewis, in contrast to Barra, has generally been described as marking the distinction between epenthetic and non-epenthetic forms with a tonal distinction (e.g. MacAulay 1979), Ó Curnáin (1990) has shown that instead Lewis appears to be at one extreme of a continuum where as one moves further north so the F0 peak is realised later in the vowel. In Lewis F0 peaks relatively late in the vowel and therefore in long vowels the peak is reached only shortly before the vowel offset, and this applies both to the long vowel in 'bò' and the diphthong in 'fiach' as well as to the vowels in 'arm'. In contrast, in Barra the F0 peak is reached somewhat earlier and is, roughly speaking, maintained at that level throughout the vowel. Thus nothing special needs to be said about the intonation pattern of epenthetic vowels, it is simply that appropriate to long vowels and diphthongs, and while important in many ways it remains a predictable secondary phenomenon.

The identity in structure of diphthongs and epenthetic sequences is apparent even in dialects where no special intonation or stress patterns have been claimed to distinguish epenthetic vowels from otherwise identical underlying vowels. In the dialect of Applecross (Ternes 1973), for example, historical diphthongs are realised synchronically as short vowel + half-long vowel sequences. Stressed vowel - epenthetic vowel sequences are realised in precisely the same way, in contrast to 'normal' VCV sequences where both vowels are short. An instrumental study of this dialect would possibly show that the intonation patterns of the two are also identical.

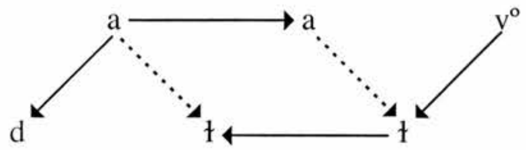
The realisation of epenthesis in Modern Irish, as we have seen, involves not a copying of the preceding vowel but simply schwa of a type not prosodically distinct from schwa elsewhere in the language. Historically speaking, there is some evidence that at an earlier stage the epenthetic vowel in Irish may have been more like that in Scottish Gaelic. For example, the accent mark (´) was used in Old Irish to mark long vowels, and also occasionally over short vowels + nonlenited sonorants e.g. 'báll' (member), and short vowels + heavy consonant groups e.g. 'árd' (high). The former of these are often realised synchronically with long vowels or diphthongs, the latter with long vowels in all dialects. Clusters which gave rise to epenthesis were also occasionally marked in the same way, hence 'árm'. In terms of poetic metrics, these two types constituted 'middle quantity' (síneadh meadhonach) (Greene 1952), so that a word class from one could rhyme with a word from the other e.g. 'árd' / 'árm'. In later, less strict metres, these could also rhyme with simple long vowels. While this is still the case in Scottish Gaelic, epenthetic sequences are no longer felt to be prosodically distinct in modern Irish.

Diachronically speaking, however, while we cannot say for certain that both Scottish and Irish Gaelic were identical in the prosody of the epenthetic vowel, it seems certain that it was in both languages prosodically distinct from non-epenthetic vowels. In terms of vowel quality, however, we cannot assume that spellings such as 'árm' suggest a pronunciation of the Scottish type, only that it suggests that such epenthetic sequences were identical (or at least very similar) in structure to underlying long vowels. Dependent vowels in closed syllables are of necessity overlapped by coda consonants, and this may cause a diminution in the overlapped vowel to the effect that it may not reach its target. The effect of this is that it may be that from the beginning the dependent (i.e. epenthetic) vowel in 'árm' was in Irish to be identified with schwa. Aside from such diminution, another factor may be the differing realisation of F0 in the two languages. Ó Cúrnáin (1990) notes that in Lewis, where V^2 is a copy of V^1 most regularly, F0 peaks at the offset of the nucleus, as noted above. In contrast, in Tiree, where the epenthetic vowel is much less commonly a copy of the preceding vowel, and where it also undergoes a degree of centralisation towards schwa, F0 peaks early in the vowel, earlier still than in Barra. If such a difference between the two languages were present at the time epenthesis first occurred, it could help to account for the different realisations in each.

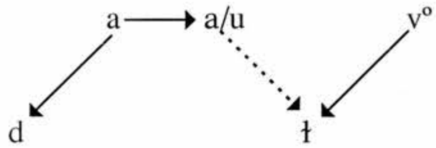
Whatever the diachronic facts behind epenthesis in Irish, synchronically we can say that in both languages the syllable structure of 'arm' contains a long vowel which coordinates with a following coda and onset. However, whereas in Scottish Gaelic the second half of this long vowel generally surfaces intact to give /a/, in Irish the dependent vowel is realised as schwa if a coda is present. The reduction of vowels to schwa is a common process in Irish, where the majority of unstressed short vowels are so reduced. When no coda is present, however, no reduction takes place, as can be seen in (13). In (13a) /daɫ:/ (blind) represents the putative Old Irish pronunciation (and also the synchronic pronunciation in most dialects of both Scottish and Irish Gaelic). In many dialects, however, such apparently short vowel + geminate sonorant sequences result in the form in (13b) where the vowel is lengthened and/or diphthongised. This reduces simply to the loss of the coda, and with no coda present there is no vowel reduction.

(13)

a) [dał:]



b) [da:ł] or [dauł]



The important point to note in terms of the phonology of Gaelic is that a single analysis can be given for the two languages, where epenthetic sequences are monosyllabic in both. The sole difference between them is that in Irish the dependent vowel reduces to schwa when a coda is present.[11] In terms of AP itself, the evidence from both Gaelic and Winnebago clearly indicates the primacy of phonological relationships over physical ones. We recall that the existence of a direct physical relationship between two segments always implies the existence of a phonological relationship between the two. Winnebago and Gaelic show that the reverse is not true, that is, an abstract, phonological relationship can hold between two segments without an accompanying physical relationship. Given the inequality between the vocalic and consonantal tiers, however, this holds true only on the consonantal tier. Consonants can drift apart because ultimately they are all bound together by their licensers on the vocalic tier.

As Steriade (1990) points out, it is the inherent physical and temporal extent of gestures which makes possible the analysis put forward here. Theories without

[11]It might even be possible to say that the dependent vowel is these instances is now underlyingly schwa. However, this would create a number of new underlying diphthongs. On the other hand, we

such physical and temporal aspects are unable to capture the relationship between the gradient nature of such consonantal drift and its resulting categorical effects. The syllable structures proposed here are able to maintain the essence of Steriade's analysis of Winnebago whilst at the same time constraining it to avoid the overgeneration of apparently phonologically relevant patterns of coordination. The same is true of Gaelic, where the same general processes as those attested in Winnebago give a similar result. As well as supporting the syllable structures proposed here, we also find straightforward explanations for the various apparently anomalous suprasegmental features of the epenthetic sequences in the two languages.

The caveat remains that the set of data tested remains small, but within that confine the set of phonological and physical relationships set out in both this and the previous chapter seem capable of providing some genuine insights into phonological and phonetic structures and the relationship between them. What remains in terms of intersegmental structures is to extend the analysis into domains such as the foot and metrical theory in general, as well as to provide analyses of syllable structure for a wider range of data. The theory, or rather theories, proposed here regarding syllable structure constrain to a greater or lesser degree the previously unbounded possibilities of gestural coordination to a set of possibilities bounded by phonological structure. At the same time, the physical properties of AP allow us to capture both the categorical aspects of syllable structure without recourse to abstract timing units, and at the same time leave us able to capture the most fine grained of gradient features. Whether or not this is an attractive property is a question which can only be answered with time.

---

need to provide an explanation as to how the weakening of the dependent vowel to schwa in epenthetic sequences is accomplished.

5.4
Constriction Degree
5.4.1
Sanskrit

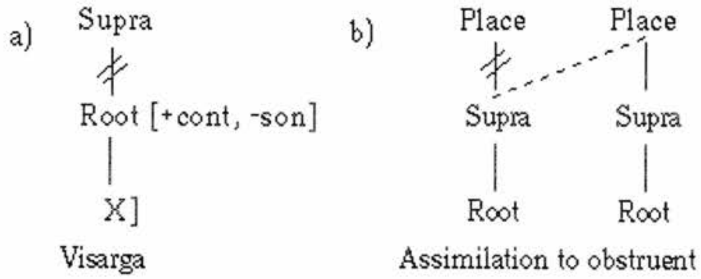
In addition to highlighting some of the perceived benefits of adopting a gestural approach, Steriade (1990) also suggested that there is a range of data which can not be adequately described in terms of gestures without fatally weakening the link between constriction degree and constriction location. These involve a number of spreading processes cited by Sagey (1982) as evidence that place and manner, in featural terms, must be independent features of any phonological approach. Padgett (1991) removes much of the difficulty in his revised geometry where [cont] is dependent on the articulator nodes, but the data in (14) from Sanskrit nevertheless remain a problem. A process known as visarga applies to word-final /s/, whereby it is debuccalised, the result being /h/ ('ḥ', however, is the traditional transliteration and will be used here, in contrast to 'h' which represents /ɦ/). Schein & Steriade (1986) note that there is in addition an optional process assimilating /s/ to a following obstruent e.g. Nala[x] kamam in (14c).[12] Schein & Steriade's rules are given in (15).

(14)

a.	Indras śuraś	'the hero'	>	Indraḥ śuraḥ	~	Indra[ś] śuraḥ
b.	tas ṣaṭ	'those six' (fem)	>	taḥ sat	~	ta[ṣ] sat
c.	Nalas kamam	'at will'	>	Nalaḥ kamam	~	Nala[x] kamam
d.	divas putras	'God's power'	>	divaḥ putraḥ	~	divaɸ putraḥ

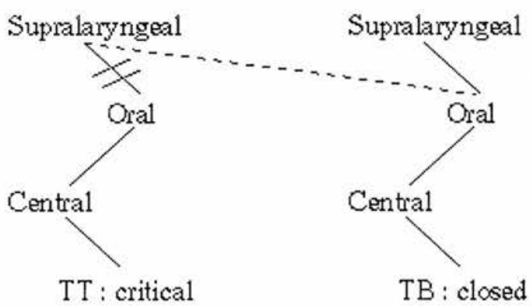
[12]Another, obligatory, rule assimilates all coronal obstruents to a following coronal stop. This does not affect the discussion in hand.

(15)



The crucial point is that assimilation involves delinking /s/ from its own place node and linking it to the place node of the following obstruent. This means that the original value of [cont] of /s/ is preserved, as can be seen in (14) where assimilation always results in a fricative. This of course poses a serious problem to a gestural analysis where place and manner are not separable, at least in terms of spreading or deleting one without the same processes applying to the other. Taking for example /s/ assimilating to /k/, /s/ consists of TT^H : **cri**, GLO : **open**, /k/ of TB^H : **clo**, GLO : **open**. In each of these gesture, place and manner are intimately bound. Spreading the oral TB gesture of /k/ would result not in Nalax kamam but *Nalak kamam (16) (assuming that assimilation takes place at the level of the oral tube). The only apparent way out of this quandary is to abandon the indissoluble link between constriction degree and constriction location and to allow assimilation to spread location only. This would involve weakening the link not only for the spreading consonant /k/ but also on the target /s/, as the CL of /s/ would also need to be deleted.

(16)



The resulting theory would be almost indistinguishable from FG, and would completely undermine the gestural basis of the AP. However, an alternative solution to the problem of visarga is possible, one in which there is no need for any weakening of the CD / CL link. While Schein & Steriade assume that assimilation bleeds visarga, there is no a priori reason why the same approach should be taken in AP. Visarga can, as noted above, be described as debuccalisation, where the supralaryngeal level of /s/ is removed. In gestural terms the TT gesture is deleted, leaving behind only the GLO gesture.[13] Importantly, this involves not only a change in the gestural content but also in the status of the gestures. For /s/, the TT gesture is a head, the GLO gesture a non-head. For /ḥ/, however, the GLO gesture is the sole head. Creation of visarga, then, moves us from a segment in which an oral gesture is the sole head to a segment in which a glottal gesture is the sole head.

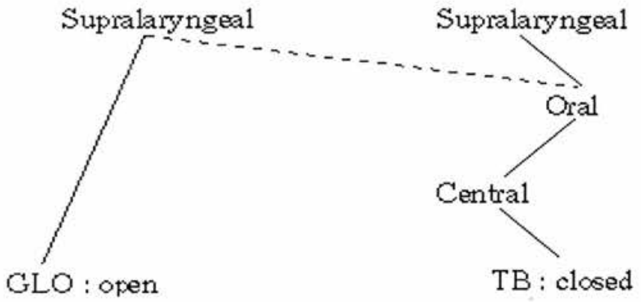
What happens if, unlike Schein & Steriade, we assume that visarga occurs without exception, i.e. assimilation does not bleed it? Clearly visarga is a weakening process of some kind, whereby the word-final position is unable to license the presence of the oral gesture for /s/.[14] Equally important, however, is the fact that the glottal gesture does not delete along with the oral gesture but instead remains and moves from non-head to head. Thus visarga is a weakening of the oral gesture - all the way to deletion - and at a same time a strengthening of the glottal gesture. Assimilation must be seen to occur in this dual environment. The oral gesture which spreads in (17) (again representing Nalas + kamam) is a head within the segment from which it spreads, and it spreads to a segment in which a GLO gesture is a head. Given the existence of both voiced and voiceless aspirated stops in Sanskrit, we might therefore expect assimilation in Nalas + kamam to result in *Nalak^h kamam, i.e. the creation of a segment with two heads. However, although Sanskrit contains both voiced and voiceless aspirated stops, the only consonants allowed occur word-finally are /k ʈ t p ŋ n m r ḥ/ (not including the fricatives found in visarga environments), i.e. word-final position does not tolerate segments with more than one

[13]Although I speak of deletion, there is of course no reason to suppose that a process of deletion is present. /s/ will have two separate forms: /ḥ/ in visarga environments, and /s/ elsewhere.

[14]The /s/ is not strictly word-final, as it is licensed by a following empty vowel.

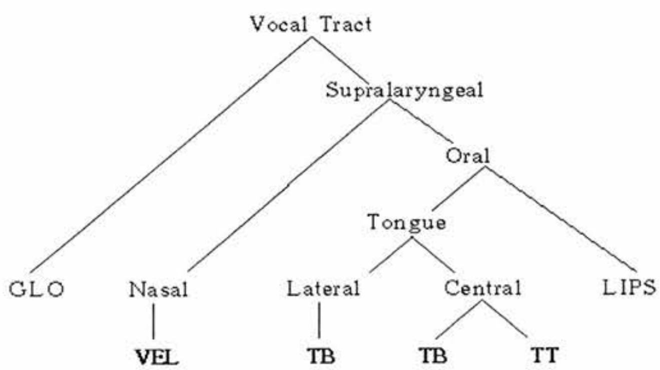
head. Following assimilation in (17), then, only one of the gestures - TB : **clo** or GLO : **open** - can be a head, but which is it?

(17)



If visarga applies without exception then clearly it is the GLO gesture which must be the head. As noted above, visarga involves not only a weakening of the oral gesture but also a strengthening of the glottal gesture. The assumption that oral gestures are always heads reflects the importance of such gestures crosslinguistically. In particular, oral gestures can be regarded as providing the articulation to which glottal gestures provide the phonatory backdrop, so that glottal gestures are the secondary modification of the primary oral gestures. Looking again at the vocal tract hierarchy in (18), we can see that it can be divided neatly into two halves at the vocal tract level, one containing the various tubes forming the supralaryngeal tube, the other containing simply the glottal tube. What would happen, however, if the dominance of these two halves were to be reversed, i.e. if the glottal gestures were instead primary, the oral gestures now being a secondary modification?

(18)



This is what seems to be happening in the assimilated forms in (14) above. Visarga first creates a segment in which a glottal opening gesture is the sole head; assimilation then adds an oral gesture but now this oral gesture is not a head but a non-head, so that assimilation in Nalas + kamam creates GLO^H : **open**, TB : **clo**. The computing of the overall value of the vocal tract for non-complex segments discussed in chapters 1-3 was based on the supralaryngeal tube containing the head. Given a segment of the form GLO : **open**, TB^H : **clo** the oral gesture would constrain the output at the vocal tract level to be silence. While the noise from the glottis would be present, and have a significant effect on the overall acoustic output, it is nevertheless the oral gesture which would dominate.

The creation of a segment GLO^H : **open**, TB : **clo**, where the GLO gesture is the sole head, means that it is the glottal tube which must now acoustically dominate. When no assimilation occurs GLO^H : **open** is the sole gesture and this results in an open vocal tract where the output at the vocal tract level is one of noise. The same would be true when other non-headed gestures are present, i.e. whatever the state of the rest of the vocal tract the overall output must be one of noise. The TB gesture is therefore constrained so that no matter its CD, it must be compatible with the demand for an output of noise at the vocal tract level. What this means is that the TB gesture will move towards its goal and will produce a CD which approaches this goal as closely as possible whilst obeying the constraint placed upon it regarding the overall value of the vocal tract. This will result in the TB producing a CD of critical constriction in line with the Glottal gesture's goal of creating noise as the overall value of the vocal tract. This is, of course, exactly what we find in (14) where assimilation produces fricatives in every case

Given this analysis, Steriade's criticism of the relationship between CD and CL is considerably weakened. The facts of Sanskrit are no longer at odds with a gestural approach, indeed this approach provides a straightforward solution which links the creation of visarga with the alternative forms containing word-final fricatives, and does so without any need for rule ordering or bleeding. Assimilation can now be seen to involve not just Place (or CL), but assimilation of an entire gesture.

This analysis assumes that there is real assimilation. However, we recall that Icelandic preaspiration is best represented not as spreading of a glottal gesture from one segment to another, but as a glottal gesture occupying a coda position. If we were to analyse visarga in the same way then we would have a process directly comparable with preaspiration, with this time an oral gesture spreading to occupy a coda position of a previous word which terminates in /h/. If this is so, then clearly only the glottal gesture can be the head, as no segment as such is created from the combination of the glottal and oral gestures, and the physical realisation will still follow that outlined in the discussion above.[15]

5.4.3
Icelandic

The question now arises as to whether this analysis of visarga is anything more than sleight of hand. The assumption must be that such a radical restructuring of the nature of headedness can not be an isolated instance. Further support comes from the spirantisation processes in Icelandic, touched upon briefly in chapter 4 and illustrated in (19). Both /pʰ/ and /kʰ/ lenite to /f/ and /x/ when followed immediately by /tʰ/ or /kʰ/, and occasionally when followed by /s/. In (19a) the addition of both the neuter inflection /tʰ/ and the preterite ending /tʰɪ/ cause the preceding stop to spirantise. In each case, the forms with stops are all preceded by long vowels (or a long diphthong in the case of [kau:pa]), and given the syllable structures proposed in chapter 4 we can say that these stops are in onset position. In contrast, the forms with fricatives are all preceded by short vowels, which would suggest that the fricatives are not in onset position but instead occupy the coda.

---

[15] Padgett (1991) goes further and suggests that no assimilation takes place at all and that fricative realisations of visarga are simply the phonetic realisation of preaspiration. Regardless of whether phonological assimilation how does or does not occur, and of how we should represent assimilation in AP, we still have to account for the surface forms.

(19)

a.

ríkur	[ri:kʏr]	'rich' masc.	ríkt	[rixt]	'rich' neut.
rakur	[ra:kʏr]	'stiff' masc.	rakt	[raxt]	'stiff' neut.
djupur	[dju:pʏr]	'deep' masc.	djupt	[djuft]	'deep' neut.
kaupa	[kau:pa]	'buy' impv.	keipti	[keiftɪ]	'buy' preterite

b.

litka [lɪ:tʰkʰa] ~ [lɪ:tka] ~ [lɪhtka] ~ [lɪθka] 'colour'

c.

bát	[pau:t]	'boat'	báts	[pau:ts]	'boat' gen. sg.
skip	[scɪ:p]	'ship'	skips	[scɪ:ps] ~ [scɪfs]	'ship' gen. sg.

In (19b) we see the four realisations of 'litka' given by Oresnik & Pétursson (1972). As noted in chapter 4, there is an alternation between forms with long and short [ɪ], an alternation which again suggests that the /tʰ/ may occupy the onset or the coda. In [lɪhtka] the stop must be analysed as a geminate, i.e. it occupies both the onset and the coda positions simultaneously, leading to shortening of the vowel. In [lɪθka], where the stop is lenited, there is again shortening of the vowel. Lenition of the labial and velar stops is relatively common in that a number of inflections create the necessary environment. The neuter and the preterite endings /tʰ/ both trigger lenition, but of course when added to a stem which itself ends in /tʰ/ the result is a geminate. However, [lɪθka] shows that lenition of /tʰ/ is also possible. The generalisation appears to be that, rather than claiming that these stops are lenited when followed by certain other stops, they are instead lenited when they are syllabified solely in the coda. The forms in (19c) add further weight to this conclusion. As 'báts' illustrates, addition of the genitive singular inflection /s/ does not lead to shortening of the vowel, and therefore the preceding stop is not syllabified into the coda. However, the alternative realisations of 'skips' shows that vowel shortening can occur. Crucially, when the vowel is shortened, and hence when the stop is syllabified into the coda, it is realised not as a stop but as a fricative.[16]

All three aspirated stops are thus realised as fricatives when they occupy the coda, so that in featural terms there is an alternation between realisations with [-cont] and [+cont]. In gestural terms it appears to be an alternation between Oral : **clo** and Oral : **cri**. Here it is clear that the featural system has a clear advantage in that there is variation within a single feature, whereas the change from a closed to a critical gesture, although moving from one degree of closure to the next, is a shift from one category to another and in that sense quite arbitrary. As Kohler (1992) points out when discussing lenition processes in AP, if this is a gradient rather than a categorical change whereby the oral gesture fails to reach its target through a process of diminution, there is seemingly no way to ensure that the now 'correct' diminished target of critical closure is instead achieved. In other words, if we simply diminish the closure gesture, how do we ensure that diminution does not continue down to an open gesture, ultimately leading to deletion of the consonant altogether? Even as a categorical process, the change from a closed to a critical gesture is entirely arbitrary and no more natural than a change to an open gesture, and is unexplanatory regarding why this change occurs only to aspirated stops, and only when in the coda.

In fact, the phenomena in (19) seem to be of essentially the same kind as seen above for Sanskrit. In Southern Icelandic, both aspirated and non-aspirated stops are found word-initially, but neutralisation of the aspirate - non-aspirate distinction occurs intervocalically in favour of the non-aspirates. In Northern Icelandic, neutralisation takes the reverse direction, in favour of the aspirates. In gestural terms, neutralisation revolves around the question of whether the glottal opening gesture of stops should be headed or non-headed in intervocalic position. The effect in Southern dialects is that intervocalically only a single headed gesture is allowed. In addition to this neutralisation, we know that geminate aspirates are realised as

---

[16]This raises all kinds of questions regarding segmental strength, as we would need to provide an answer as to how it is that /s/ is able to govern the apparently stronger, i.e. less sonorous, stop. I leave this matter open for further research.

preaspirates, in other words the GLO gesture of these stops, not the oral gesture, occupies the coda position. We therefore have two processes affecting stops, both involving the glottal opening gesture. When non-geminate aspirated stops are syllabified into the coda, I suggest that the GLO gesture again dominates, but that in this case it does so by being the sole head. The result of course would be identical to that in visarga, with the underlying stops being realised as fricatives. Again there is no change to the gestural content of the segments, only a change in the phonological relationships between them. We also avoid an arbitrary and unexplanatory change in the CD of the oral gesture, and provide an explanation as to why this process occurs only to the aspirated stops, linking it directly to the internal headed structure of the segment and the behaviour elsewhere in the phonology of the GLO gesture.

Here, then, are two instances in which the expected internal structures of segments within AP are reversed. In both cases we have evidence that outside of the instances of creation of spirants, the glottis can be a headed gesture. In Sanskrit, there is no doubt that /s/ lenites to /h/, and in Icelandic there is no doubt that geminates are realised as sequences of /h/ followed by a stop. In AP terms, both of these involve headed GLO gestures. The vocal tract hierarchy itself is divided in two, with the vocal tract tube being a combination of the supralaryngeal and glottal tubes. Viewed from this perspective, it is not so surprising that, as it were, the balance of power between the two tubes should be open to change

However, while the data from Sanskrit involved processes of assimilation, that from Icelandic did not. The implication of lenition in Icelandic is thus that similar processes in other languages should be describable in the same way. That is, if we can identify a language as having a synchronic process (or processes) of lenition, there should be at least some cases in which lenition can be described not as a categorical change in the value of an oral gesture, but as a result of the increased prominence of a glottal gesture. Lenition in Old Irish, Welsh, and Spanish I suggest can be described in this way. The situation in each case is complex, and each are deserving of an in depth analysis. Here, however, I wish simply to put forward some tentative analsyses, none of which are wholly complete, concentrating on the segmental changes involved in order to illustrate the advantages or otherwise of the approach.

5.4.3
Old Irish

The data in (20) are from Old Irish.[17] Lenition here affects both sets of stops, voiceless and voiced, where they become the corresponding fricatives. Lenition also affects the fricative /s/ which becomes /h/ and in addition, the labial nasal /m/ is lenited to /ṽ/.[18] In featural terms we would again classify the changes to the stops as one involving [cont], whether involving changing an existing value i.e. [-cont] to [+cont], or some kind of fill-in process (see Ní Chiosáin (1991) for a discussion of various accounts of lenition in modern Irish). The same could apply to the nasal /m/. For /s/, though, there is a different type of change, the same process which we have already seen in Sanskrit whereby the oral gesture deletes and the glottal opening gesture becomes the sole head.

(20)

pobal	/pobəɫ/	'people'	a phobal	/ə fobəɫ/	'his people'
tuath	/tuaθ/	'people'	a thuath	/ə θuaθ/	'his people'
capall	/kapəɫ/	'horse'	a chapall	/ə xapəɫ/	'his horse'
bó	/bo:/	'cow'	a bó	/ə vo:/	'his cow'
dam	/dãv/	'ox'	a dam	/ə ðãv/	'his ox'
gó	/go:/	'falsehood'	a gó	/ə ɣo:/	'his falsehood'
mac(c)	/mak/	'son'	a mac	/ə ṽak/	'his son'
salm	/saɫm/	'psalm'	a salm	/ə haɫm/	'his psalm'

[17]The representation is simplified somewhat, as the Old Irish consonantal system had a two way distinction, whereby every consonant was either palatalised or velarised, and these distinctions had an effect on the lenited reflexes. These contrasts are not shown here.

[18]/f/ is also lenited to zero and clearly in a full description this would need to be accounted for. Old Irish, like its modern Gaelic descendants, also had a process of lenition which affected the sonorants. This was not a process of spirantisation, however, and does not affect the argument in hand here.

The fact that /s/ is so affected suggests that the glottis might also have a role to play in the lenition of the other obstruents. For /p t k/ this would involve a change from Oral^{H} : **clo**, GLO : **open** to Oral : **clo**, GLO^{H} : **open**, where headedness is reversed. For /b d g/ a similar process applies but here the headed GLO gesture has a CD of critical i.e. Oral : **clo**, GLO^{H} : **cri**. I shall deal with the expected consequences of this shortly, but one problem immediately occurs in that for these voiced stops there is, according to the conception of neutral settings proposed by Browman & Goldstein (1989), no GLO gesture present. The change here would therefore involve the addition of such a gesture, i.e. Oral^{H} : **clo** becoming Oral : **clo**, GLO^{H} : **cri**. The alternative would be for /b d g/ to contain an active as opposed to neutral GLO gesture. The consequences of such a move need to be investigated, but at present I assume that lenition in both cases involves an alternation between forms with an Oral head and forms with a GLO head.

What though would be the result of making GLO : **cri** the sole head? The percolation principles discussed in chapter 1 (recalling however that these are based on the tacit assumption that the supralaryngeal gestures are the heads) suggest that if the supralaryngeal tube is open and the glottal tube is critical, the result at the vocal tract is resonance, a value encompassing anything from a nasal to a vowel. Following lenition we have GLO^{H} : **cri** gestures, alongside (for /b d g/) Oral closure gestures. The requirement at the vocal tract level is, I suggest, still one of resonance due simply to the presence of voicing, but can be either noisy resonance, i.e. a voiced fricative, or non-noisy, i.e. a semivowel, the precise realisation varying from language to language. In Old Irish, lenition of the voiced stops is presumed to have resulted in the creation of voiced fricatives, given the reflexes in the modern dialects. There is some variation regarding the realisation, in Modern Irish, of the lenited reflex of both /b/ and /m/ in terms of presence or absence of frication. For example, in the Irish of Cois Fhairrge (de Bhaldraithe 1945) lenited /b/ and /m/ (ignoring nasality) are realised as [w] word-initially before a vowel, but as [β] or [v] elsewhere. Similar variation is found in other dialects, where the distinction between fricative and approximant realisations is clearly not categorical but gradient and allophonic.[19] This is not to say that fricatives and semivowels should not be distinguished, only that they should not be so distinguished when they result from this type of lenition.

As for /m/, aside from the question of why this sonorant and no other is prone to spirantisation (as opposed to other types of lenition),[20] this approach has one clear advantage, and that is that by avoiding a change within the CD of the oral gesture we also avoid creating a nasal segment with an Oral CD of critical, a segment type which is at the least very rare and possibly non-existent (Padgett 1991). A change in the CD of the LIPS gesture from closure to critical would not necessarily lead to noise, as the lowered velum would vent the airflow. In such conditions the percolation principles predict that, unless greater articulatory force than usual is present, the result at the supralaryngeal level and thus at the vocal tract level will be an open tube, not a critical one. As we have seen above, lenition of /m/ can result in a fricative in all the modern Gaelic dialects, and presumably this was also the case in Old Irish. By modelling lenition not as a direct change to the Oral gestures but as an increase in the prominence of the glottis, these problems are avoided. The presence of a $\mathrm{GLO^H}$: **cri** gesture demands, in certain conditions, that the overall output will be one of noise despite the presence of a lowered velum.

A further point may be made regarding the role of the glottis in segmental weakening in Gaelic, and that is the realisation of the inherited voiceless - voiced contrast of the stops in Old Irish as voiceless aspirated - voiceless unaspirated in most Scottish Gaelic dialects. In these dialects the contrast is one of $\mathrm{GLO^H}$: **open** vs GLO : **open**, with an increase in prominence of the glottis of a different kind. In

---

[19]Ní Chiosáin (1991) suggests a process of glide formation involving both a change from [-son, +cons] to [+son, -cons], and a demotion of the features [Labial] and [Dorsal] to [+round] and [+high] respectively. This complex change is unnecessary given the approach taken here.

[20]A full description of the consonantal phonology would clearly be required before we could provide a complete account of lenition in either Old Irish or its modern descendants. One point regarding /m/ may be relevant, however. The labial closing gesture for /m/ in Scottish Gaelic appears to be considerably weaker than that for its English equivalent (Mike Broe, personal communication) in that in Gaelic the target is simply closure while in English the target is a point considerably beyond closure, showing a large amount of lip compression. This comparative weakness may be connected with its role in the lenition system, though this would of course need substantiating.

addition, in many Scottish Gaelic dialects the aspirated stops may, in certain environments, be realised as preaspirates of a type comparable to those of Icelandic. Whatever the historical origin of this development, preaspiration again clearly involves an increase in the prominence of the role of the glottis, suggesting at least a connection between the development of lenition and preaspiration.[21] Borgstrøm's claim that 'there is nothing in the Common Gaelic or Irish systems to explain why the Scottish Gaelic sound-shift has taken place' (Borgstrøm 1940, p. 219) is at the very least questionable.

Finally, while certain dialects realise preaspiration in a way similar to the corresponding sounds in Icelandic, i.e. as [hC] (or [^{h}C] in the case of Lewis), others have developed an oral fricative realisation in certain environments. In Barra, for example, preaspiration before (non-palatal) dental and labial stops surfaces as [ht hp], but the preaspirated velar stop is [xk].[22] Ní Chasaide & Ó Dochartaigh (1984) suggest that the relatively slow movement of the tongue body in velar stops may have led to the period of preaspiration being overlapped by the oral closure gesture, somehow leading to [xk]. We can compare this with the realisation of prenasalised stops, where the nasal portion always cooccurs with an oral portion. If this were to be the case, the result in AP would be predicted to be identical to that found in assimilation to visarga discussed above, e.g. Nala[x] kamam. This is, of course, exactly what we find.

[21]There is disagreement as to whether preaspiration is a feature developed due to historical contact with Old Norse as opposed to a purely a native development. The extraneous origin seems unlikely, especially since there is in fact no evidence that Old Norse itself had developed preaspiration during the period when the two languages were in close contact. The evidence that preaspiration was present in Old Norse at this time rests almost exclusively on the fact that Scottish Gaelic also has preaspiration (Marstrander 1932) and that it must have borrowed it from Old Norse, a circular argument.

[22]A number of dialects have [xt xp xk], with a velar fricative in all realisations. It seems clear, however, that the velar fricative developed first before [k] and its appearance before the other stops is a later development (Clements 1983)

5.4.4

Welsh

The role of the glottis in consonantal weakening is perhaps clearer in the Aspirate Mutation (AM) of Welsh (21). In both the literary language and the spoken language this involves an alternation between voiceless stops and the corresponding fricatives (21a). While not reflected in the literary language, in many dialects AM also affects the nasals /m n/, which are realised as voiceless or aspirated nasals [m̥ n̥] ~ [m̥h n̥h], forms identical with the realisation of the Nasal Mutation of voiceless stops seen in chapter 3. These appear to be, in featural terms, two entirely separate changes. On the one hand we have a change in the value of [cont] from [-cont] in the case of the stops, to [+cont] for the corresponding fricatives. On the other hand, for the nasals we have presumably the addition of either [-voice] or [+SG], or both. An analysis where a single change instead affected both sets of consonants would clearly be preferable.

(21)

a.

cath	[ka:θ]	'cat'	ei chath	[i xa:θ]	'her cat'
pen	[pɛn]	'head'	ei phen	[i fɛn]	'her head'
tad	[ta:d]	'father'	ei thad	[i θa:d]	'her father'

b.

mam	[mam]	'mother'	ei mham	[i m̥am]	'her mother'
nan	[nain]	'grandmother'	ei nhain	[i n̥ain]	'her grandmother'

It is debatable whether Welsh has a voiced - voiceless contrast in its stops, or a voiceless aspirated - voiceless unaspirated contrast. Whatever the case, AM for the stops would, in the model proposed here, again be analysed as an increase in the prominence of the glottis so that a GLO^{H} : **open** gesture became the sole head. These segments would then be realised as the corresponding fricatives. The effect of AM on the nasals would be slightly different. If these are to be analysed as [m̥h n̥h] i.e.

as containing two heads, rather than as [m̥ n̥] with a single head, then the GLO^{H} : **open** gesture would be one of two heads. If on the other hand the variation seen in the realisation of these voiceless nasals as to whether or not they show complete overlap of the various gestures - whether caused by AM of the nasals, or Nasal Mutation of the stops - is purely gradient, then GLO^{H} : **open** would be the sole head in all cases. Again, a fuller examination of Welsh, taking in an analysis of the consonantal system and the associated mutations as a whole, would be needed before a definitive answer could be provided. What is important is that while the effect of AM on the two sets of consonants is markedly different in terms of the physical realisations, the underlying mutation may be identical in each case.

5.4.5
Spanish

Finally, I turn to the question of spirantisation of the voiced stops in Spanish (I will refer to this below as lenition, for reasons which will become clear). This has been well described in a number of works, e.g. Mascaró (1984), Harris (1984), Padgett (1991), and the generalisation seems to be that voiced stops /b d g/ lenite to /β ð ɣ/ when they follow a [+cont] segment. The data in (22), taken from Padgett (1991), shows that lenition occurs when the stops follow a vowel, glide, fricative or liquid. Utterance initially and following nasals there is no change to the stops. Padgett, in common with Mascaró and Harris, assumes that the process involved is one of rightwards spreading of [+cont], this spreading from the preceding continuants in (22a-d). Note that this is a feature changing rule, deleting as it does the preexisting [-cont] node of the underlying stop. As nasals are [-cont] lenition will clearly not occur following them, and the fact that stops do not lenite utterance initially now follows from the fact that there is no preceding [+cont] node available for spreading. In terms of targets, [+cont] will spread to segments which are specified [+voice]. Voiceless stops are thus not affected, and the only remaining voiceless segments are themselves already [+cont]. Note that this view of lenition would appear to predict that nasals should also be targets, but Padgett, along with

Branstine (1991), assumes that lenition of nasals is blocked by a marking condition against nasal fricatives *[+nas, +cons, +cont].

(22)

Relevant segments are in bold

a. (vowels)		b. (glides)	
ca**β**ello	'hair'	cai**ɣ**a	'he falls' (subj.)
a**βð**ica	'he abdicates'	dew**ð**a	'doubt'
la**ð**o	'side'		
la**ɣ**o	'lake'		

c. (liquids)		d. (fricatives)	
cur**β**a	'curve'	dis**ɣ**usto	'trouble'
cal**β**o	'bald'	a**βð**ica	'he abdicates'
pur**ɣ**a	'purge'	des**ð**e	'since'
al**ɣ**o	'something'	a**ðβ**erso	'adverse'

[-cont] elsewhere

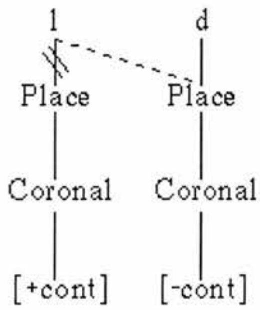
e. (after pause)		f. (after nasal)	g. /d/ after /l/
bueno	'good'	e**ŋg**ano	ca**l̪d**o
gana	'he pays'	ho**mb**ro	
dedo	'finger'	ho**n̪d**o	
gato	'cat'	cua**n̪d**o	

It is clear that spreading must be iterative, as we can see in (22d) where [+cont] spreads first from the vowel to the following coronal stop in [aðβerso], and continues to spread to the labial stop, halting only when another [+cont] segment is reached. This appears to present a considerable problem, as the possibility of spreading [+cont] does not exist in AP, and the only thing which would seem to unite the various triggers - vowel, glide, liquid and fricative - is that at the vocal tract level they are all *not* closed. All that would be available for spreading, then, would be this negative value of the overall output of the vocal tract, which seems an unattractive option to say the least. However, the picture given here of Spanish is somewhat idealised, and there are a number of problems with Padgett's analysis, both in terms of its theory-internal predictions and in terms of its ability or otherwise to handle dialects (which are in fact by far the majority) where the facts are not so neat.

Given Padgett's analysis and the fact that both /b/ and /g/ spirantise following /l/, it is clear that /l/ must be [+cont]. Why, then, does /d/ also not spirantise in the same environment? In (22g) we see that although /l/ is generally alveolar, preceding /d/ i.e. [d̪] it is dental [l̪], indicating that there is some linked structure, specifically the [Place] node of the stop spreading rightwards to the sonorant, whose own [Place] node is delinked in the process (23). This spreading occurs before [+cont] spreading. Given Padgett's geometry, where [cont] is a dependent of the [Place] node (though we recall that [cont] is actually a dependent of each individual articulator, these articulators themselves being dependent on [Place]), assimilation will remove the [+cont] node. As /l/ is no longer [+cont] spreading will not occur, hence no lenition.[23] Similarly, as no assimilation occurs before /b/ or /g/, /l/ will spread its [+cont] node and cause lenition. Unfortunately, this creates a segment /l/ which can be either [+cont] or [-cont] with no phonetic effects whatsoever. While this may be justified in terms of the phonological analysis, it greatly weakens the link between phonological features and their phonetic realisation, as there is nothing in the geometry of /l/ which would predict that the value of [cont] is irrelevant in terms of the phonetics.

[23]Lenition is also blocked as a result of linked structure, as I discuss below.

(23)



There is another type of weakening process cited by Amastae (1986), though not mentioned by Padgett, to which voiced stops are prone (24). In (24a) it appears as if there is voicing assimilation between the voiced stops and the following voiceless fricatives with the effect that the stop is devoiced. In other words, there is an apparent process of [-cont] spreading from voiceless stops to voiced stops, and this must occur before lenition, thereby bleeding it. The forms in (24b) however rule out the possibility of this being due to voicing assimilation, as here devoicing occurs before nasals which are themselves voiced. The only other option is to class this as a process of segmental weakening occurring in a particular syllabic environment, i.e. lenition, and this is what Amastae himself suggests. Given that voiced stops are prone to this process of lenition, it may be that spirantisation of the stops is also due to lenition rather than [+cont] spread, despite Padgett's rejection of this option. The advantage would of course be that if neither devoicing nor spirantisation are caused by spreading or assimilation of any kind we remove the problem of [+cont] spreading.

(24)

a.

ab̥surd ad̥junto

b.

ab̥negar ad̥mitir ignorar sub̥marino

By way of contrast to the elegant picture given in (22), Amastae (1986) reports a far more varied picture. In the dialects he examined, non-lenition was regular in two environments, utterance-initially and following nasals, appearing to provide some support for Padgett's analysis. Elsewhere, however, there was considerable variation as to whether /b d g/ were realised as stops or fricatives; in many environments in which we would expect to find fricatives, Amastae found that instead stops were very often, and in some cases much more often, found. Some of the variation was simply due to the consonant involved, so that /d/ was more likely than either /b/ or /g/ to be realised as lenited in all environments. Much of the variation, however, seemed to involve syllable structure and in particular stress. All stops were less likely to lenite when they occupied onsets as opposed to codas. Intervocalically, if the preceding vowel was stressed then all three stops were far likelier to lenite than if stress fell on the following vowel.[24]

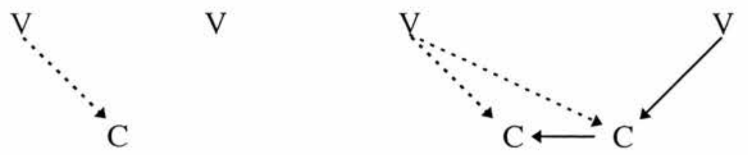
These findings highlight the fact that, at least in certain environments, the presence or absence of lenition is variable. The factors involved may be to some extent sociolinguistic, but clearly syllable structure, and almost certainly foot structure, play an important part. The temptation is therefore to analyse lenition as a process affecting voiced stops in syllabically weak positions, though the identity of these weak positions is still to be determined. In terms of the syllable structure developed so far (and assuming Spanish is like Icelandic and not Italian), the two primary environments of lenition can be described as in (25). In (25a), lenition occurs intervocalically, while in (25b), the stop licenses a coda, whether a glide, liquid, fricative or stop. In both environments, however, the stop is directly coordinated with the preceding vowel and in both environments the stop lenites. Utterance initially this is, of course, not the case, and we can therefore provisionally state that voiced stops will lenite when they are indirectly licensed by both a following and a preceding vowel.

---

[24]The text actually seems to suggest that the reverse is true. However, Amstae's data show that the picture given here is accurate, i.e. spirants are more commonly found when the preceding vowel is stressed.

(25)

a) b)



As we saw above, /l/ assimilates in place to a following /d/ and in so doing blocks lenition. The /d/ however is still coordinated with the preceding vowel, as in (25b), so why does lenition not occur? The casual speech process in Cuban Spanish known as Havana liquid assimilation, discussed in chapter 3, results in realisations such as 'cu**bb**a' for standard 'cur**β**a', where in gestural terms /r/ and /l/ are completely overlapped by a following consonant. We saw that for e.g. 'cu**bb**a' < 'curba' the entire gesture for /b/ is doubly linked to the preceding vowel, completely overlapping the preceding /r/ and resulting in a geminate stop. This double linking blocks lenition.

Padgett explains the failure of lenition following Havana liquid assimilation by reference to the Linking Constraint (26), given the representation of lenition in (27). As (27) does not mention double linking, lenition will not target assimilated forms created by Havana liquid assimilation. This also explains why, once /l/ in /l + d/ sequences becomes [-cont], the [+cont] value of the preceding vowel does not spread first to /l/ and then to /d/.[25] This in fact highlights a general difficulty with the linking constraint, to the effect that although [+cont] spreading must be iterative, the linking constraint actually rules this out, a problem to which Padgett offers no solution. Of course, if lenition is not caused by [+cont] spread but is due to segmental weakening, this is no longer a problem.

---

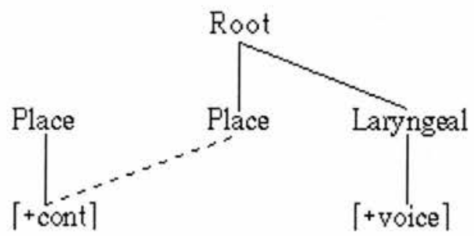
[25] In many dialects, e.g. El Salvador, Nicaragua (Canfield 1981), lenition is blocked when a voiced stop licenses a coda of any kind, so that lenition will only occur intervocalically as in (22a).

(26)

<u>Linking Constraint</u> (Hayes 1986)

Association lines in structural descriptions are interpreted as exhaustive

(27)



As we saw in (25), in the case of both an intervocalic voiced stop and of coda + voiced stop the stop coordinates with both the preceding vowel and the preceding consonant and is therefore lenited. In the case of both Havana liquid assimilation and /l + d/ sequences, however, the stop, or part of it, occupies the coda position. We could express this fact in a number of ways, such as invoking a version of Hayes' (1986) Linking Constraint, or Steriade & Schein's (1986) Uniform Applicability Condition. Alternatively, we could follow Scobbie (1991) and simply block lenition by having a condition on its output, i.e. geminate spirants are disallowed. I will not attempt to decide between these competing theories, but will instead informally note that if a voiced stop, in whole or in part, occupies both onset and coda then lenition does not occur. A full answer of the lenition environments must surely wait until a full account is given of what constitutes a weak environment in keeping with the findings of Amastae.

Linked structure (or double coordination) might also be invoked to explain the non-occurrence of lenition following a homorganic nasal (22f). As an argument against this, Padgett cites dialects where the nasal is not homorganic to the stop but is instead realised as a velar nasal whatever the nature of the following stop, giving realisations such as 'cua[ŋ]do' instead of cua[n̪]do. As there is no linked structure here, Padgett rules out the possibility of linked structure blocking lenition in nasal + stop environments. The occurrence of lenition in e.g. 'desde' but never in 'cuando', along with the assumption that the stop is in the same position in the syllable in each case, therefore appears to rule out the possibility of an analysis involving lenition in weak syllable (or foot) positions.

However, Padgett fails to take fully into account the nature of the [n] ~ [ŋ] alternation in 'cuando'. Spanish contains nasals distinguished for three places of articulation, namely palatal, dental and labial. These occur freely in onset position, but preconsonantally and word-finally there is complete neutralisation within the series. In most dialects (except those, of course, with realisations such as cuaŋdo) nasals are homorganic to the following consonant no matter the identity of that consonant i.e. stop, fricative or other (Harris 1969). In certain circumstances homorganicity is not an option i.e. word-finally, and then there is neutralisation of the place contrast and in standard Spanish this neutralisation is realised as the coronal nasal [n]. In the dialects cited by Padgett, neutralisation is instead realised as the velar nasal [ŋ], so that [ŋ] is found both word-finally and in word-internal coda, hence realisations such as [k^{w}aŋdo].

It is thus only those dialects which have the velar nasal in all neutralised contexts which cause problems for a blocking approach.[26] This difficulty is illusory, however. In the syllable structure developed here, [ŋ] word-finally is in fact licensed by a following empty vowel, and we can therefore make the generalisation that an empty vowel in Spanish can only license a single CL for nasals, either TT for [n] or TB for [ŋ], depending on dialect. In those dialects where nasals in the coda i.e. word-internally are also realised as [ŋ], the simplest solution is that these are also directly licensed by a following empty vowel, i.e. they are no longer in codas but are now underlyingly in onset position. This gives us a representation of 'cuando' as [k^{w}aŋv^{o}do]'. The fact that /d/ does not lenite is now simply due to the fact that it does not fulfil the criteria, that is, it is not indirectly governed by a preceding vowel.

---

[26] In other dialects, [ŋ] is found only word-finally, with /n/ as the neutral form word-internally. The arguments here apply equally to these dialects.

Again, the situation in individual dialects is clearly very complex, and a full analysis would need to take into account the variation in where and when lenition occurs as highlighted by Amastae. A full analysis is also needed of the implications of linked structures in AP and why it is that they block processes which otherwise would be expected to apply, but this is not a problem for AP alone. The structures developed here shed light on a number of aspects of Spanish syllable structure, but at the same time Spanish highlights a number of areas where AP needs to be developed. The tentative account given here of the relationship between syllable structure and lenition has one clear advantage from the point of view of AP in that it avoids the description of lenition as involving [+cont] spreading, a process which is impossible in AP. Instead, lenition must be seen as a process taking place solely within a segment, and by doing so we connect the change from /b d g/ to /β ð ɣ/ with a number of other apparently unrelated processes in Spanish.

As we have already seen, in certain environments the voiced stops may be devoiced, e.g. ab̥negar. In gestural terms this involves the addition of a non-headed glottal opening gesture, and the addition of such a gesture in this environment strongly suggests that spirantisation be viewed as one of two lenition processes affecting voiced stops. If this is so, spirantisation would also involve the addition of a GLO gesture, but in this case it would be GLO^H : **cri**. If this were again the sole head, spirantisation would automatically produce [β ð ɣ] with no need for spreading of any kind. Devoicing and spirantisation could then both be classed under the general heading of lenition. However, unlike in the languages examined so far, there is no evidence that GLO gestures can be heads elsewhere in the stop series, and this may perhaps be seen as a weakness despite the existence of the process of devoicing. Evidence of the ability of GLO gestures to be heads must be found outside of the stops.

Stops are not the only segments in Spanish which can be lenited. In the majority of dialects /s/ is lenited to /h/ word-finally and in the coda, and again this must be analysed as an increase in the prominence of the glottis.[27] By analysing lenition of the voiced stops as involving the glottis we provide a link between all three phenomena - debuccalisation, devoicing, and spirantisation - rather than having them each as separate developments. More suggestive still is the behaviour of /f/ and /x/ in a number of South American dialects as investigated in Lipski (1995). In standard Spanish /f/ is realised as a voiceless labio-dental fricative, and /x/ as a voiceless velar or uvular fricative. In non-standard dialects /f/ may instead be realised as a voiceless bilabial fricative [ɸ], and in many of the same dialects /x/ may be realised as [h]. Lipski notes that in dialects with these realisations, the two consonants show a number of interesting interchanges. Let me note here that although Lipski speaks of /x/ being 'realised' as [h] and /f/ being realised as [ɸ], these realisations are in fact exceptionless (not counting the variations discussed below) and not allophonic variants occurring in certain circumstances. In these dialects there is no segment /x/, only /h/, no segment /f/ only /ɸ/. Following Lipski I will refer to these segments as [h] and [ɸ] respectively.

In the dialect of the Peruvian Amazon, Lipski notes that historical [ɸ] is often replaced by [h] before [+round] segments, i.e. before syllabic or semivocalic /u/, and less commonly before /o/ (28a) ('j' represents [h]). The forms on the left, representing the standard orthography, reflect the fact that more prestigious forms of the language retain the labial fricative. Nevertheless, realisations with [h] are in fact widespread and not confined to the area examined by Lipski. Malmberg (1950), discussing the dialect of Buenos Aires and its surrounding area, notes that before rounded vowels and semivowels [ɸ] and [h] are in free variation. In the dialect examined by Lipski, however, the forms with [h] are now fixed i.e. there is no longer any synchronic alternation between forms with [ɸ] and those with [h]. Lipski characterises [ɸ] as [+round, +cont] sharing the specification of [+round] with

[27]In fact, Canfield (1981) notes that lenited /s/ in both Uruguay and Argentina may have an an oral articulation in certain environments, giving the example of 'busca' [buxka] where the TB gesture of the following /k/ appears to spread to the headed GLO gesture in a manner reminiscent of visarga and of preaspiration in Barra Gaelic. Canfield notes that the realisation of /s/ as [x] is especially common when /k/ follows, but (Malmberg 1950) suggests that the precise realisation of /s/ is dependent upon

rounded vowels in contrast to other labials which are specified as [+labial]. The change from [ϕ] to [h] then involves deletion of [+round].[28]

(28)

a.

difunto	>	dijunto	[dihunto]
fondo	>	jondo	[hondo]
fue	>	jue	[hwe]
fuego	>	juego	[hweɣo]
fumar	>	jumar	[humar]

b.

familia	~	juamilia	[ϕamilia]	~	[hwamilia]
facíil	~	juacíl	[ϕasil]	~	[hwasil]
fecha	~	juecha	[ϕetʃa]	~	[hwetʃa]
filo	~	juilo	[ϕilo]	~	[hwilo]

c.

Juan	~	Fan	[hwan]	~	[ϕan]
juez	~	fez	[hwes]	~	[ϕes]
juicio	~	ficio	[hwisjo]	~	[ϕisjo]

The forms in (28b,c) have a more restricted distribution, apparently occurring in those dialects in which the process in (28a) is no longer an active synchronic one. In (28b) we see forms in which the historic /f/ shows alternation between [ϕ] and [hw] before unrounded vowels. (28c) shows the reverse process in which historic /xw/ shows alternation between [hw] and [ϕ] in the same context. Lipski discusses the alternations in connection with the nature and interaction of [round] and [labial]

---

the quality of the preceding vowel rather than on the following consonant. Further investigation might prove interesting.

[28] I will not discuss the merits or otherwise of representing /h/ as [+cont]. See the discussion in chapter 3.

in Feature Geometry. He suggests that the forms in (28b,c) are no longer distinct, and that the (partial) underlying representation is as in (29) where there is a floating [+round] node. This floating node is free to link leftwards to the consonant to give [ɸ], or rightwards to the vowel to give [w].[29]

(29)

```
Root
 |
[+cont]

          [+round]
```

Looking initially at the (diachronic) change in (28a), at first blush the bilabial fricative [ɸ] would appear to have a gestural content of LIPS^{H} : **cri**, GLO : **open**, while [h] would be GLO^{H} : **open**. The change from [ɸ] to [h] would then involve a deletion of the LIPS gesture to leave behind only the GLO gesture, which would then become the sole head in a process parallel to the lenition of /s/ to [h]. While this appears reasonable, the connection between [ɸ] and [hw] is now opaque. On the one hand we have a single segment with an internal structure of LIPS^{H} : **cri**, GLO : **open**, and on the other we have a sequence of two segments with a structure of GLO^{H} : **open** for [h] followed by LIPS^{H} : **open** for [w].[30] While we can see that GLO and LIPS gestures are present in each case, that seems to be as much as we can say. Lipski presents a convincing argument that, at least in featural terms, [ɸ] and [hw] contain the same segmental content, differing only in whether they are realised as a single segment or as two separate segments. It appears that in AP, however, such an analysis is not possible.

Following Lipski, let us assume that we have the two possible partial representations in (30). In (30a) the single segment contains two gestures, a glottal opening gesture, known to be present due to the voiceless nature of [ɸ] and a LIPS

[29]The remainder of the content of [w] would be added by default.

gesture of uncertain CD. Which of these gestures is a head is as yet uncertain. In (30b) the status of these gestures is clearer. As both are separate segments which contain only a single gesture, both the GLO and the LIPS gestures are heads. The fact that the LIPS gesture is realised as [w] allows us to identify its CD as **open**. Again following Lipski, let us assume that when there is a single segment the content of the gestures does not change. This allows us to fill in the CD of the LIPS gesture of (30a) as **open**, identical to the gesture in (30b). How, then, can the combination of GLO : **open** and LIPS : **open** result in [ɸ]? The straightforward answer is that in (30a) it is the GLO gesture which is the head, i.e. the gestural content of [ɸ] is not *LIPS^{H} : **cri**, GLO : **open** at all but rather GLO^{H} : **open**, LIPS : **open** (30c). The alternation between [ɸ] and [h] now follows the main line of Lipski's analysis. Synchronically, there is free variation as to whether the combination of these two gestures forms a single segment, corresponding to historical /f/, or whether they instead form a branching onset, corresponding to historical /xw/.

(30)

a)	b)		c)
C	C	C	C
GLO : **open**	GLO^{H} : **open**	LIPS^{H} : **open**	GLO^{H} : **open**
LIPS : **?**			LIPS : **open**

The importance of this for our analysis of lenition processes in Spanish is clear. The fricative realisations of the voiced stops, namely /β ð ɣ/, are analysed as containing a GLO gesture as the sole head. This analysis of [ɸ] shows that they are not the only such segments in the language, at least in the dialect under consideration. Alone, these two facts may not be considered conclusive, but, taken together with the fact that the role of the glottis appears to be prominent in both the lenition of /s/ to /h/ and the devoicing of the voiced stops to /p t k/, the case as a whole is at least persuasive. The [ɸ] ~ [h] alternation is distinct from the other

---

[30] I assume here, like Lipski, that /w/ consists of a headed LIPS gesture. The behaviour of the tongue

processes in that it seems to involve a lexical aspect absent from lenition per se. Rather than put forward a phonological process converting one into the other, it seems that the two consonants, at least in the environments in (28) and perhaps elsewhere, are simply no longer distinguished in this dialect. This is further evident in the fact that [ɸ] in clusters can also be realised as [h] e.g. 'fruta' ~ 'jruta'. The absence of [w] is due to general phonotactics of Spanish - [hw] is not possible before /u/.

These facts of Spanish are not alone, however, but instead form part of a larger pattern involving similar phenomena in Sanskrit, Icelandic, Gaelic and Welsh. These changes might be modelled as following a sonority chain of /p/ ~ /b/ ~ /f/ ~ /v/, involving changes in two unconnected areas, namely the laryngeal node and [cont]. Further changes to /h/ may or may not be incorporated into this picture. As we have seen in Spanish, these changes can be analysed as having no connection whatsoever, involving as they do such different processes. Looked at from the AP perspective, however, we instead find a unifying theme in the role of the glottis.

This is not to claim that we have provided a complete picture of the processes involved, far from it. In each of the languages discussed in this section there remain a number of questions unanswered, and in each case we can only talk of possibilities and point out suggestive similarities. While we can recognise the role of the glottis in each case, both the environment of the changes and the changes themselves differ from language to language. The facts of Spanish highlight this problem very well. The role of syllable (and possibly higher) structure in triggering lenition of Spanish voiced stops is, I believe, clear in as much as it plays a part, but the precise nature of the role of syllable structure would need to be determined by a more detailed examination of Spanish in general. Different dialects behave in different ways, however, a fact which makes it difficult, if not impossible, to label certain structures as weak and others as strong.

---

is handled by the neutral settings of the articulators as discussed in chapter 3.

In many ways the theory developed here is markedly different from competing theories, and these differences spring from the fundamental basis of incorporating gestures into models of phonological structure. These gestures bring with them a number of properties which immediately distinguish them from the features, atoms etc. of other theories, properties which prompt a reassessment of both segmental and syllabic structure At the same time, the basic framework which has been constructed, whether in terms of internal segment structure or intrasegmental coordination, must be seen in the light of the discussion in chapter 2. In other words, what may look like a radical departure from previous thought will turn out upon closer inspection to be no more than a linear development from what preceded. This is as true of AP as it is of other theories and should be taken as a point in its favour. While there should not be a blind incorporation of the basic frameworks of other theories, neither should we reject them out of hand. Aside from attempting to ensure that all phonological units are built up from gestures, there is little limit to the kind of structures which we could build. The strength, or lack of it, of the structures proposed here rests solely upon their ability to explain and predict. It is upon this basis that further development must proceed.

Chapter 6
Further Developments

Articulatory Phonology presents us with a new view of phonology and phonetics and the manner in which the two are related, stretching the barrier between the two fields to the point in which it disappears. The question of whether or not the abandonment of the division between phonology and phonetics is achievable or even desirable is obviously a matter of some debate (see Coleman 1992 for an opposing point of view), but in practical terms we may find an answer to this question once we are able to fully determine how well or badly AP fares in its attempts to straddle the phonology-phonetics barrier. At present, however, the resolution of this matter lies some way in the future as despite the relative successes of AP in dealing with a number of different phenomena, there remains a great many areas in which it may be said to fail as a theory of phonology.

The basic units of AP - gestures - are well defined physical objects which interact in a clearly defined anatomical hierarchy. In addition, the physical interactions of these objects, both in terms of their physical production and overlap as well as the audio-acoustic output, is fully implemented in a working model, and to this extent any claims that we make regarding the structures of speech within the framework of AP are directly testable. At the same time, we have seen that although a number of gradient phenomena receive simple and illuminating explanations in a gestural framework, much of this is undermined by the general absence of any formal framework which would allow us to describe categorical structures. Processes such as the apparent deletion of the /t/ in `perfec**t** memory' or the assimilation of the nasal in 'bea**m**bag' have been convincingly analysed as resulting from changes in the temporal ordering of the component gestures, and this is straightforwardly illustrated in the gestural scores of each phrase by increasing the overlap between gestures. These gestural scores, however, are claimed to contain categorical structures which might be equated with segments, e.g. the tongue tip and glottal gestures for /t/, but there is nothing within the gestural score which is able to distinguish between the interaction of these gestures in creating /t/ and the interaction of the same two gestures in a sequence such as /t + h/. It is not enough to say that categorical structures are present, we must show that they are.

The lack of any explicit means of capturing categorical information makes it difficult, if not impossible, to describe processes such as nasal assimilation in Bantu languages which, while similar to the casual speech assimilation in 'bea**m**bag', is quite different in that assimilation in the Bantu languages involves a discrete, categorical change from one segment to another. Assimilation in 'bea**m**bag', on the other hand, is only apparent as the underlying coronal nasal may in fact still be present. The distinction between the two is fundamental, and AP's lack of ability to distinguish between the two has been heavily criticised. It should be clear, however, that this is far from being a fundamental flaw in AP, and instead it is best characterised as something which is yet to be added to the theory.

Leaving aside the question of how AP should account for categorical structures, we must also identify precisely what these categorical structures are, and in attempting to do this there are two readily identifiable areas which other phonological theories have consistently identified as essential to a proper description of speech, namely the segment and the syllable. The evidence for both of these levels of organisation is strong and indeed their existence is implicit in many of the structures of AP (Browman & Goldstein 1988, 1989), but, as we have seen, however strong the evidence for these levels there is no obvious way in which they should be described. The result is, of course, that different phonological theories have developed which each provide different formalisations of these levels, these different formalisations being more or less compatible with each other depending on the particular topic under discussion (Roca 1994). There is nothing to prevent us adapting any one of these different formalisms to the needs of AP, but in fact the strong resemblance of the Vocal Tract Hierarchy to the numerous hierarchies of Feature Geometry has led to the partial adoption of the formalisation of FG in Browman & Goldstein (1989). However, the structures proposed there have a number of problems, some simply mirroring similar problems with the formalism of

FG, others stemming from the adaptation of the structures of FG to a set of articulatory primitives rather than to an abstract feature set. We are then left with the decision as to whether we attempt to find a solution to these problems and present a theory of AP within a framework similar to FG, an attractive proposition given the increasingly close relationship between the Vocal Tract Hierarchy and the geometries of FG (Padgett 1991). Alternatively we might instead attempt to describe the phonological relationships between gestures in quite a different way. It is this latter course which has been explored here.

One of the main obstacles facing us in attempting to account for phonological regularitics in a gestural framework is the difficulty in providing a non-arbitrary account of the linear ordering between gestures, both at the level of the segment and of the syllable, this ordering implicit in the gestural score. The overlap between two or more gestures presents us with a continuum which we can divide in an infinite number of ways, and clearly overlap alone generates a far greater number of relationships between gestures than phonologically necessary or useful. Overlap without ordering would result in unintelligibility. What is needed is a means of deriving the ordering between gestures in such a way so as to avoid any arbitrary partitioning of the overlap continuum, whilst not robbing the power of the theory to continue to capture fine grained gradient distinctions.

This is precisely what the head-dependent relations outlined here are designed to achieve, ensuring that AP is able to provide a full account of categorical processes whilst continuing to distinguish between the categorical and the gradient. In other words, while the /m/ in 'seve**m** plus' and that in /hem/ may appear in the stream of speech to be identical, the head-dependent structures provide us with a simple means of distinguishing between the two. Without such a means, speakers would be unable to access the correct lexical form of 'seven' with a final /n/. The phonological relationships between gestures within a segment are thus defined directly in terms of the temporal relationships which hold between them, so that where two gestures are offset with respect to each other or alternatively completely overlap each other, this is directly reflected in the phonological structure of the segment.

The advantage of describing both gradient and categorical information with the same basic primitives is, I suggest, seen in many areas. For example, in phenomena such as nasal spreading in Gurarani and French, there are in each case categorical processes at work but the structures of AP developed here drastically reduce the number of phonological rules necessary by arguing that many apparently categorical effects are in fact gradient. Similarly, the contrast described by Connell (1994) in the result of nasal assimilation to /gb/ - i.e. /ŋgb/ or /ŋmgb/ - is shown not to be categorical but gradient. These analyses are very much in the spirit already developed in AP whereby a number of casual speech phenomena are shown to be better analysed in terms of gradient overlap of gestures, and the introduction of a framework for representing discrete structures allows us to profitably explore how the two domains interact. It must be said, however, that the theory developed here is obviously not the only possible framework for AP, rather the measure of its usefulness lying in how well it describes the data. Given this, in order to provide a more thorough test of the framework the range of segment types so far described must be extended, as must the range of phonological phenomena, and this extension will inevitably bring with it changes to the framework, a process which is to be welcomed.

The same problems associated with the lack of categorical structures is also apparent in the AP's approach to syllables, though there is a more explicit picture available of the coordination which exists between gestures at this level (e.g. Browman & Goldstein 1989, 1990b, Byrd 1996). Here, too, I propose the introduction of head-dependent structures, where the phonological relationships are now between segments rather than gestures. While the languages investigated share much in common, there are also major between them which make crucial use of the variation possible in coordination. For processes such as epenthesis in Winnebag and Scottish Gaelic, explanations which do not make use of at least some physical element comparable to a gestural approach face a number of very real problems. Again, in order to substantiate the claims made here for syllable structure we need to widen the range of phenomena described, and given the direct reliance of the phonological structures on variation in the coordination between segments, it is essential that we make more explicit the account of the physical coordination between gestures by implementation in a physical model. Pen and paper only take us so far.

This is not to negate the value of phonological speculation in the absence of a physical model, as implementation of every nuance of every individual theory is clearly a daunting, if not impossible, task. It is more in the nature of acknowledging that the physical nature of AP, and the claims made regarding the role of categorical and gradient structures, means that a full assessment of the theory developed here can only be made by explicitly testing its claims in a physical way. In the absence of such a physical implementation we cannot refer to the head-dependent structures developed here as a model of speech, only as a theory. Nevertheless, even given this restriction, these structures provide us with a means of constraining the physical relationships between gestures in a way which answers the criticisms of e.g. Clements (1992) whilst still maintaining the flexibility needed in a theory which claims to breach the phonology-phonetics barrier. This, at least, seems a goal worth striving for.

Bibliography

Amastae, J (1986) A Syllable-based Analysis of Spanish Spirantisation. In Jaeggli, O & C Silva-Corvalan (eds.) Studies in Romance Linguistics. Dordrecht:Foris

Anderson, S (1976) On the Description of Multiply Articulated Consonants. Journal of Phonetics **4**. 17-27

Anderson, J & C Ewen (1987) Principles of Dependency Phonology. Cambridge. Cambridge University Press

Archangeli, D (1983) The Root CV-Template as a Property of the Affix: Evidence from Yawelmani. Natural Language and Linguistic Theory **1**. 347-384

Archangeli, D (1985a) Yokuts Harmony: Evidence for Coplanar Representation in Non-linear Phonology. Linguistic Inquiry **16**. 335-372

Archangeli, D (1985b) CV-skeleton or X-skeleton: The Turkish Evidence. Studies in the Linguistic Sciences **15**. 1-10

Arnason (1980) Quantity in Historical Phonology. Cambridge. Cambridge University Press

Arnason (1986) The Segmental and Suprasegmental Status of Preaspiration in Modern Icelandic. Nordic Journal of Linguistics **9**. 1-23

Ball, M (1984) Phonetics for Phonology. In Ball, M & G E Jones (eds.) Welsh Phonology. Cardiff. University of Wales Press

Barry, M C (1985) A palatographic study of connected processes. Cambridge Papers in Phonetics and Experimental Linguistics **4**. 1-16

Beckman, M, J Edwards & J Fletcher (1992) Prosodic Structure and Tempo in a Sonority Model of Articulatory Dynamics. In Docherty, G J & D R Ladd (eds.)

Bell-Berti, F (1980) Velo-pharyngeal function: A spatial-temporal model. In N J Lass (ed.) Speech and Language: Advances in Basic Research and Practice. Volume 4. New York. Academic Press

Bendor-Samuel, J T (1960) Some Problems of Segmentation in the Phonological Analysis of Terena. Word **16**. 348-355

Benguerel, AP & T U McFadden (1989) The effect of coarticulation on the role of transitions in vowel perceptions. Phonetica **46**. 80-96

de Bhaldraithe, T (1945) The Irish of Cois Fhairrge, Co. Galway: A Phonetic Study. Dublin Institute for Advanced Studies. Dublin

Bird, S (1990) Constraint-based Phonology. PhD dissertation. University of Edinburgh

Bird, S & E Klein (1990) Phonological Events. Journal of Linguistics **26**. 33-56

Borgstrøm, C H (1940) The dialects of the Outer Hebrides. A linguistic survey of the Gaelic dialects of Scotland. Volume 1. Norsk tidskrift för sprogvidenskap.

Borgstrøm, C H (1941) The dialects of Skye and Rossshire. Norsk tidskrift för sprogvidenskap. Supplementary volume 2

Borgstrom, C H (1942) Scottish Gaelic and Early History of Irish. In Scottish Gaelic Studies **5**. 33-44

Bosch, A (1988) VC syllable structure in Scottish Gaelic: some implications for syllable theory. Unpublished paper. University of Kentucky

Bosch, A (1995) A Gestural Analysis of Epenthesis in Scottish Gaelic. Unpublished paper. University of Kentucky

Boyce, S E & R A Krakow & F Bell-Berti & C E Gelfer (1990) Converging Sources of Evidence for Dissecting Movements into Core Gestures. Journal of Phonetics **18**. 173-188

Branstine, Z (1991) Stop/Spirant Alternations in Spanish : On the Representation of Contrast. In Studies in the Linguistic Sciences **21**. 1-24

Brentari, D, & A Bosch (1990) The Mora: Autosegment or Syllable Constituent. In Ziolkowski et al.

Brockhaus, W (1991) Final Devoicing and Neutralisation. In UCL Working Papers in Linguistics **3**.

Browman, C P & L Goldstein (1985) Dynamic modelling of phonetic structure. In V Fromkin (ed.) Phonetic Linguistics: Essays in honor of Peter Ladefoged. New York: Academic Press

Browman, C P & L Goldstein (1986) Towards an articulatory phonology. Phonology Yearbook **3**. 219-252

Browman, C P & L Goldstein (1988) Some notes on syllable structure in articulatory phonology. Phonetica **45**. 140-155

Browman, C P & L Goldstein (1989) Articulatory Gestures as phonological units. Phonology **6**. 201-256

Browman, C P & L Goldstein (1990a) Tiers in Articulatory Phonology, with some implications for casual speech. In Kingston & Beckman (1990)

Browman, C P & L Goldstein (1990b) Gestural Specification Using Dynamically-Defined Articulatory Structures. Journal of Phonetics **18**. 299-320

Browman, C P & L Goldstein (1991) Distinctiveness, phonological processes and historical change. In I Mattingly & M Studdert-Kennedy (eds.) Modularity and the motor theory of speech perception. Erlbaum. Hillsdale

Browman, C P & L Goldstein (1992a) Articulatory Phonology: An Overview. Phonetica **49**. 155-180

Browman, C P & L Goldstein (1992b) Response to Commentaries. Phonetica **49**. 222-234

Byrd, D (1992) Perception of Assimilation in Consonant Clusters: A Gestural Model. Phonetica **49**. 1-24

Byrd, D (1996) A phase window framework for articulatory timing. Phonology **13**. 139-169

Cairns, C E & M H Feinstein (1982) Markedness and the Theory of Syllable Structure. Linguistic Inquiry **13**. 193-225

Canfield, D L (1981) Spanish Pronunciation in the Americas. Chicago. University of Chicago

Catford, J C (1977) Fundamental problems in phonetics. Bloomington. Indiana University Press

Charette, M (1991) Conditions on phonological government. Cambridge University Press. Cambridge

Chomsky, N & M Halle (1968) The Sound Pattern of English. New York. Harper and Row.

Clements, G N (1985) The Geometry of Phonological Features. Phonology Yearbook **2**. 225-252

Clements, G N (1992) Phonological Primes : Features or Gestures. Phonetica **49**. 155-180

Clements, G N (1994) Place of articulation in sonsonants & vowels: a unified theory. In B Laks & A Rialland (eds.) L'architecture et la géometrie des representation phonologiques. Paris. Editions du CNRS

Clements, G N & S Keyser (1983) CV Phonology : A Generative Theory of the Syllable. Massachusetts. MIT Press

Clements, R D (1983) Gaelic : Preaspiration. In D S Thomson (ed.) The Companion to Gaelic Scotland. 104-105. Oxford. Blackwell

Cohn, A (1990) Phonetic and Phonological Rules of Nasalization. UCLA Working Papers in Linguistics **76**. 1-224

Coleman, J (1992) The phonetic interpretation of headed phonological structures containing overlapping constituents. Phonology **9**. 1-44

Coleman, J & J Local (1991) The 'No Crossing Constraint' in Autosegmental Phonology. Linguistics & Philosophy **14**. 295-330

Connell, B (1994) The structure of labial-velar stops. Journal of Phonetics **4**. 441-476

Dixit, R P (1989) Glottal gestures in Hindi Plosives. Journal of Phonetics **17**. 213-237

Dixon, R M W (1980) The Languages of Australia. Cambridge. Cambridge University Press

Docherty, G J & D R Ladd (eds.) (1992) Papers in Laboratory Phonology, II: Gesture, Segment, Prosody. Cambridge. Cambridge University Press

Dorian, N (1978) East Sutherland Gaelic. Dublin Institute for Advance Studies.

Dunn, M H (1990) A Phonetic Study of Syllable Structure in Finnish & Italian. In Ziolkowski et al.

Durand, J (1987) On the Phonological Status of Glides : The Evidence from Malay. In J Anderson & J Durand (eds.) Explorations in Dependency Phonology

Elcock, W D (1975) The Romance Languages. Faber. London

Ewen, C (1982) The Internal Structure of Complex Segments. In van der Hulst & Smith (eds.)

Farnetani & Kori (1990) Rhythmic Structure in Italian Noun Phrases: A Study of Vowel Durations. Phonetica **47**. 50-65

Feinstein, M H (1979) Prenasalization and Syllable Structure. Linguistic Inquiry 10. 245-278

Foster, M L (1969) The Tarascan Language. University of California Publications in Linguistics **56**. Berkeley & Los Angeles

Fowler, C A (1983) Converging sources of evidence on spoken and perceived rhythms of speech: cyclic production of vowels in sequences of monosyllable stress feet. Journal of Experimental Psychology : General **112**. 386-412

Garnes, S (1974) Quantity in Icelandic: Production and Perception. PhD dissertation. Ann Arbor

Gibb, L (1993) Domains in phonology: with evidence from Icelandic, Finnish and Kikuyu. PhD dissertation. University of Edinburgh

Givón, T (1974) Rule un-ordering: generalization and degeneralization in phonology. CLS. Papers from the parasession on Natural Phonology.

Goldsmith (1990) Autosegmental and Metrical Phonology. Basil Blackwell. Cambridge. MA

Goldstein, L (1990) On articulatory binding: comments on Kingston's paper. In Kingston & Beckman (1990)

Greene (1952) Middle Quantity in Irish. Ériu **16**. 212-218

Gregores, E & J A Suàrez (1967) A Description of Colloquial Guaraní. Mouton. The Hague

Harris, J (1969) Spanish Phonology. Cambridge. MA. MIT Press.

Harris, J (1990) Segmental complexity and phonological government. Phonology **7**. 255-300

Harris (1992) Licensing Inheritance. UCL Working Papers in Linguistics **4**. 359-406

Hayes, B (1986) Inalterability in CV Phonology. Language **62**. 321-351

Hayes, B (1989) Compensatory Lengthening in Moraic Phonology. Linguistic Inquiry **20**. 253-306

Herbert, R K (1986) Language Universals, Markedness Theory, and Natural Phonetic Processes. Berlin: Mouton de Gruyter

Hermans, B (1984) The relation between Aspiration and Preaspiration in Icelandic. In van der Hulst & Smith (eds.) (1984) Features, Segmental Structure and Harmony Processes. 2 volumes. Dordrecht : Foris

Hoffman, C (1963) A Grammar of the Margi Language. London. Oxford University Press.

Holmes, J N & I G Mattingly & J N Shearme (1964) Speech Synthesis by Rule. Language and Speech **7**. 127-143

Holmer, N (1938) The Gaelic of Kintyre. Dublin Institute for Advanced Studies

Hualde, J G (1987) On Basque Affricates. West Coast Conference on Formal Linguistics **6**. 77-89

Hualde, J G (1991) Basque Phonology. London. Routledge van der Hulst, H & C Ewen (eds.) (1982) The Structure of Phonological Representations. Dordrecht: Foris

Hyman, L (1985) a theory of phonological weight. Dordrecht:Foris

Iverson, G & C A Kesterson (1989) Foot and Syllable Structure in Modern Icelandic. Norse Journal of Linguistics **12**. 13-45

Iverson, G (1989) On the Category Supralarjyngeal. Phonology **6**. 285-303

Jakobson, R & G Fant & M Halle (1951) Preliminaries to Speech Analysis. Cambridge MA. MIT Press.

de Jong, K (1991) An Articulatory Study of Consonant-Induced Vowel Duration Changes in English. Phonetica **48**. 1-17

Jónsson, J G (1994) The feature [asp] in Icelandic Phonology. Studia Linguistica **47**. 28-45

Kaye, J (1990) `Coda' Licensing. Phonology **7**. 301-330

Kaye, J & J Lowenstamm & J-R Vergnaud (1990) The internal structure of phonological segments. Phonology **7**. 305-329

Keating, PA (1988) Underspecification in Phonetics. Phonology **5**. 275-292

Kenstowicz, M & C Kisseberth (1979) Generative Phonology: Description and Theory. San Diego. Academic Press

Kingston, J (1985) The phonetics and phonology of the timing of oral and glottal events. PhD dissertation. University of California, Berkeley

Kingston, J (1990) Articulatory Binding. In Kingston & Beckman (1990).

Kingston, J & M Beckman (1990) Papers in laboratory phonology 1: between the grammar and the physics of speech. Cambridge. University of Cambridge
Kiparsky (1985)

Kingston, J & A Cohen (1992) Extending Articulatory Phonology. Phonetica **49**. 194-204

Kim, C W (1970) A theory of aspiration. Phonetica **21**. 107-116

Kiparsky, P (1985) Some Consequences of Lexical Phonology. Phonology **2**. 85-138

Kohler, K (1976) Die Instabilität wortfinaler Alveolarplosive im Deutschen: eine elektropalatographische Untersuchung. Phonetica **33**. 1-30

Kohler, K (1992) Gestural Reorganisation in Connected Speech : A Functional Viewpoint on 'Articulatory Phonology'. Phonetica **49**. 205-211

Kornfilt, J (1986) Stem-penultimate empty C's, compensatory lengthening and vowel epenthesis in Turkish. In Wetzels, L & E Sezer (eds.) (1986)

Krakow, R A (1989) The articulatory organization of syllables: a kinematic analysis of labial and velar gestures. PhD dissertation. Yale University

Ladefoged, P (1990) On dividing phonetics and phonology: commentaries on the papers by Clements and by Browman & Goldstein. In Kingston & Beckman (eds.)

Ladefoged, P & I Maddieson (1996) The Sounds of the World's Languages. Blackwell. Oxford

Lass, R (1976) English Phonology and Phonological Theory: Synchronic and Diachronic Studies. Cambridge. Cambridge University Press

Laver, J (1994) Principles of Phonetics. Cambridge. Cambridge University Press

Lipski, J M (1995) [round] and [labial] in Spanish and the 'free-form' syllable. Linguistics **33**. 283-304

Local, J (1991) Modelling Assimilation in Nonsegmental, Rule-Free Synthesis. In Docherty & Ladd (eds.)

Locke, J L (1983) Phonological Acquisition and Change. New York: Academic Press

Locke, J L (1986) The Linguistic significance of babbling. In Lindblom, B & R Zetterström 'Precursors of Early Speech'. MacMillan. Basingstoke

Löfqvist, A & H Yoshioka (1981) Interarticulator Programming in Obstruent Production. Phonetica **38**. 21-34

Lombardi, L (1990) The Non-linear Organization of Affricates. Natural Language and Linguistic Theory **8**. 375-425

Loporcaro, M (1988) History and Geography of Raddoppiamento Fonosintattico : Remarks on the Evolution of a Phonological Rule. In Bertinetto, P M & M Loporcaro (eds) (1986) Certamen Phonologicum : Papers from the 1987 Cortona Phonology Meeting

MacAulay, D (1979) Some functional and distributional aspects of intonation in Scottish Gaelic: A Preliminary Study of Tones. In D P O Baoill (ed.) Papers in Celtic phonology: proceedings of the Celtic phonology conference held at NUU, 28 June - 1 July, 1977. Occasional Papers in Linguistics and Language Learning. Coleraine. Northern Ireland

McCarthy, J (1979) Formal Problems in Semitic Phonology and Morphology. PhD dissertation. MIT, Cambridge. MA

McCarthy, J (1988) Feature Geometry and Dependency: A Review. Phonetica **43**. 84-108

McCarthy, J & A Prince (1986) Prosodic Morphology. Forthcoming. MIT Press

Maddieson, I (1984) Patterns of Sounds. Cambridge. Cambridge University Press

Maddieson, I (1993) Splitting the Mora. In UCLA Working Papers in Phonetics **83**

Maddieson, I & P Ladefoged (1989) Multiply Articulated Segments and the Feature Hierarchy. UCLA Working Papers in Phonetics **72**. 116-138

Malmberg, B (1950) Etudes sur la phonetique de l'espagnol parlé en Argentine. Etudes romanes de Lund 10

Marstrander, C (1932) Okklusiver og substrater. Norsk tidskrift för sprogvidenskap **V**. 258-314

Mascarò, J (1984) Continuant Spreading in Basque, Catalan, and Spanish. In Aronoff, M & R T Oehrle (eds.) (1984) Language Sound Structure. MIT Press. Cambridge. MA

Mattingly, I (1990) The Global Character of Phonetic Gestures. Journal of Phonetics **18**. 445-52

Mhac an Fhailigh, É (1980) The Irish of Erris, Co. Mayo. Dublin Institute for Advanced Studies.

Miner, K L (1979) Dorsey's Law in Winnebago-Chiwere and Winnebago Accent. International Journal of American Linguistics **45**. 25-33

Moll, K L & R G Daniloff (1971) Investigation of the timing of velar movements during speech. Journal of the Acoustic Society of America **50**. 678-684

Murray, R W & T Venneman (1983) Sound Change and Syllable Structure in Germanic Phonology. Language **59**. 514-528

Ní Chasaide, A & C Ó Dochartaigh (1984) Some Durational Aspects of Preaspiration. In Occasional Papers in Linguistics and Language Learning. Coleraine, Northern Ireland.

Ní Chiosáin, M (1991) Topics in the Phonology of Irish. PhD dissertation. University of Massachussetts

Ni Chiosáin, M (1994) Irish Palatalisation and the Representation of Place Features. Phonology **11**

Ó Baoill, D (1980) Preaspiration, Epenthesis and Vowel Lengthening: Interrelated and of Similar Origin. Celtica **13**. 79-108

Ó Curnáin (1990) Aspects of Gaelic Word Prosody. Unpublished MSc thesis. University of Edinburgh

Ó Dochartaigh, C (1981) Vowel Strengthening in Gaelic. Scottish Gaelic Studies **13**. 219-240

Ó Murchú, M (1989) East Perthshire Gaelic. Dublin Institute for Advanced Studies

Oftedal, M (1956) The Gaelic of Leurbost, Isle of Lewis. A Linguistic Survey of the Gaelic Dialects of Scotland. Vol. **3**. Norsk tidskrift för sprogvidenskap. Supplementary volume 4

Ohala, J J (1979) Universals of labial-velars and de Saussure's chess theory. Proceedings of the 9th International Congress of Phonetic Sciences, Volume 2. Copenhagen Institute of Phonetics

Ohala, J J (1983) The origin of sound patterns in vocal tract constraints. In P MacNeilage (ed.) The Production of Speech. New York. Springer-Verlag

Ohala, J J (1990) The phonetics and phonology of aspects of assimilation. In Kingston & Beckman (1990)

Ohala, M & J J Ohala (1991) Nasal epenthesis in Hindi. Phonetica **48**. 207-220

Ohman, S E G (1966) Coarticulation in VCV utterances: Spectrographic Measurements. Journal of the Acoustic Society of America **39**. 151-168

Oller, D K & R E Eilers (1988) Similarity of babbling in Spanish- and English-learning babies. Journal of Child Language **9**. 565-577

Oresnik, J (1972) On the Epenthesis Rule in Modern Icelandic. Arkiv för Nordisk Filologi **87**. 1-32

Oresnik, J & M Pètursson (1977) Quantity in Modern Icelandic. Arkiv för Nordisk Filologi **92**. 155-171

Padgett, J (1991) Stricture in Feature Geometry. PhD dissertation. University of Massachusetts

Padgett, J (1994) Stricture and Linguistic Theory. Natural Language and Linguistic Theory **12**. 465-513

Pètursson, M (1974) Les articulations de l'islandais à la lumiere de la radiocinematographie. Paris.

Piggott, G L (1988) The Parameters of Nasaization. McGill Working Papers in Linguistics **5**. 128-177

Piggot, G L (1992) Variability in Feature Dependency: The Case of Nasality. Natural Language and Linguistic Theory **10**. 33-77

Roca, I (1994) Generative Phonology. Routledge. London

Rosentall, S (1989) The Representation of Prenasalised Consonants. Proceedings of the 7th West Coast Conference of Formal Linguistics.

Rudin, C (1980) Phonetic Evidence for a Phonological Rule: G-deletion in Turkish. In Research in Phonetics. Bloomington. Department of Linguistics. Indiana University. Report No 1. 218-232

Sagey, E (1986) The Representation of Features and Relations in Non-Linear Phonology. PhD dissertation. MIT, Cambridge. MA

Saltzman, E & J A S Kelso (1987) Skilled actions: a task dynamic approach. Psychological Review **94**. 84-106

Schein, B & S Steriade (1986) On Geminates. Linguistic Inquiry **17**. 691-744

Scobbie, J (1991) Attribute-Value Phonology. PhD dissertation. University of Edinburgh

Selkirk, E (1991) A Two Root Theory of Length. In E Dunlap & J Padgett (eds.) University of Massachusetts Occasional Papers in Linguistics **14**.

Sezer, E (1986) An Autosegmental Analysis of Compensatory Lengthening in Turkish. In Wetzels, L & E Sezer (eds.) (1986)

Shuken, C (1980) An instrumental investigation of some Scottish Gaelic Consonants. PhD dissertation. University of Edinburgh

Silverman, D & J Jun (1994) Aerodynamic Evidence for Articulatory Overlap in Korean. Phonetica **51**. 210-220

Smith, C L (1995) Prosodic patterns in the coordination of vowel and consonant gestures. In Connell, B & A Arvaniti (eds.) (1995) Phonology and Phonetic Evidence: Papers in Laboratory Phonology IV. Cambridge University Press. Cambridge

Steriade, D (1982) Greek Prosodies and the Nature of Syllabification. PhD dissertation. MIT. Cambridge. MA

Steriade, D (1990) Gestures And Autosegments : comments on Browman & Goldstein's 'Gestures in articulatory phonology'. In Kingston & Beckman (1990)

Steriade, D (1992) Segments, Contours and Clusters. In Proceedings of the XVth International Congress of Linguists. Les Presses De L'Université Laval. Sainte-Foy

Stevens, K N (1972) The quantal nature of speech.: evidence from articulatory acoustic data. In E E David & P B Denes (eds.) Human Communications: a Unified View. New York: McGraw-Hill. 51-66

Stevens, K N (1989) On the quantal nature of speech. Journal of Phonetics **17**. 3-45

Studdert-Kennedy, M (1987) The phoneme as a perceptuomotor structure. In A Allport, D MacKay, W Prinz & E scheerer (eds.) Language Perception and Production. New York: Garland

Summers, W van (1987) Effects of stress and final consonant voicing on vowel production: Articulaotry and Acoustic Analysis. Journal of the Acoustic Society of America **82**. 847-863

Ternes, E (1973) The phonemic analysis of Scottish Gaelic: based on the dialect of Applecross, Rossshire. Hamburg: Buske

Thràinsson, H (1978) On the phonology of Icelandic preaspiration. In Nordic Journal of Linguistics **1**. 3-54

Trubetzkoy, N S (1969) Principles of Phonology. Translated by C A M Baltaxe. Berkely and Los Angeles. University of California lPress

Vaissière, J (1988) Prediction of Velum Movement from Phonological Specifications. Phonetica **45**. 122-139

Venneman, T (1972) Sound change and markedness theory: on the history of the German consonant system. Linguistic Change **105**. 230-274

Weijer, J van de (1993) The Manner-Place dependency in complex segments. Linguistics **31**. 87-110

Wetzels, L & E Sezer (1986) Studies in Compensatory Lengthening. Dordrecht:Foris

Ziolkowski, M & M Noske & K Deaton (1990) CLS **26**. Papers from the 26th Regional Meeting of the Chicago Linguistic Society. Volume 2. The Parasession on the Syllable in Phonetics & Phonology.

Zsiga, E & D Byrd (1990) Acoustic evidence of overlap in consonant sequences. 118th Meeting of the Acoustical Society of America